Sgríbhinní Institiúid Árd-Leighinn
Bhaile Átha Cliath. Sraith A Uimh. 1
Communications of the Dublin Institute for
Advanced Studies. Series A No. 1

QUANTUM ELECTRODYNAMICS

by

PAUL A. M. DIRAC

DUBLIN
THE DUBLIN INSTITUTE FOR ADVANCED STUDIES
64-5 MERRION SQUARE
1943

(Reprinted 1960)

QUANTUM ELECTRODYNAMICS.

By P. A. M. DIRAC.

THE present lectures, like those of Eddington, are concerned with unifying relativity and quantum theory, but they approach the question from a different point of view. Eddingon's method is first to get the physical ideas clear and then gradually to build up a mathematical scheme. The present method is just the opposite—first to set up a mathematical scheme and then to try to get its physical interpretation. This method ought to be the easier one, because it takes the easier task first. It should be easier to get the mathematical scheme, because for this purpose the number of things one has to choose between is small. The scheme, to be acceptable, must be neat and beautiful, and the number of such schemes which pure mathematics can provide is very limited. When the mathematical scheme is decided on, the subsequent task of finding its physical interpretation is rendered easier by the existence of many points of contact between the mathematics and the physics.

We shall be concerned with the problem of the interaction of a number of charged particles with the electromagnetic field, and our object will be to get a method for calculating experimental results, such as the probability of certain particles, which are incident in a certain way, being scattered in a certain way, with perhaps the emission and absorption of certain photons. The method has to be in conformity with the general principles of physics, such as the conservation laws and Lorentz invariance, and with the laws of the quantum theory so far as they are understood.

The line of attack will be to start from the mathematical scheme of the quantum theory as it is known for the non-relativistic domain and to try to extend it to make it relativistic. We shall find that there is one natural way of doing this. But when we come to consider the physical interpretation of the mathematical scheme, we shall find that the natural extension of the non-relativistic theory leads to the appearance of negative energies and probabilities. It becomes necessary to make some new assumptions, which result in the interpretation being somewhat less direct than one could wish for, but which is all the same in agreement with all essential requirements.

§1. *The Bracket Notation.*

The most convenient notation to use in quantum mechanics for general theoretical purposes in which one wants to be able to pass freely from one representation to another is the author's bracket notation.[1] The vectors corresponding to the quantum states are denoted by symbols $|\ \rangle$, consisting of a vertical line and an incompleted bracket. If one wants to give a vector a label, a say, one inserts it in the middle, thus $|a\rangle$. Each of these vectors has a conjugate imaginary vector, corresponding to the same quantum state, which is written $\langle\ |$ with the same label. Thus, in our example, $|a\rangle$ has the conjugate imaginary $\langle a|$.

A vector $\langle b|$ and a vector $|a\rangle$ have a scalar product which is a number, and which is written as a symbolic product $\langle b|a\rangle$. (For brevity one does not put two vertical lines). The following axioms hold :—

$$\langle a|b\rangle = \overline{\langle b|a\rangle}, \tag{1·1}$$

the bar denoting the conjugate complex number, and

$$\langle a|a\rangle > 0. \tag{1·2}$$

One can operate on a $|\ \rangle$ with a linear operator, the result being another $|\ \rangle$. If the linear operator ξ operates on the vector $|a\rangle$, the result is written as a symbolic product $\xi|a\rangle$. Similarly, one can operate on a $\langle\ |$ with a linear operator, the result being another $\langle\ |$, and if the linear operator ξ operates on the vector $\langle b|$, the result is written as a symbolic product $\langle b|\xi$.

A vector $|a\rangle$ and a vector $\langle b|$ have a symbolic product $|a\rangle\langle b|$, which is a linear operator. If this linear operator operates on $|c\rangle$, the result is $|a\rangle\langle b|c\rangle$, a numerical multiple of $|a\rangle$, and if it operates on $\langle c|$, the result is $\langle c|a\rangle\langle b|$, a numerical multiple of $\langle b|$.

In all our symbolic multiplication processes the associative axiom of multiplication holds, but not, in general, the commutative one. Of course when a numerical factor appears in a product, such as the $\langle b|c\rangle$ and the $\langle c|a\rangle$ in the preceding paragraph, it can be commuted with any other factor.

As an example to familiarize us with this notation, let us study the conditions for a set of vectors $\langle r|$ to form a base of a representation in ordinary non-relativistic quantum mechanics. The conditions are :—
The vectors must form a *complete* set, i.e. it must be possible to expand any $\langle\ |$ in the form

$$\langle\ | = \Sigma_r c_r \langle r|. \tag{1·3}$$

---

[1] Dirac, Proc. Camb. Philos. Soc. 35, p. 416 (1939).

They must satisfy the orthogonal normalizing equation

$$\langle r \mid s \rangle = \delta_{rs}, \tag{1·4}$$

s being another label for one of them.

These two conditions are necessary and sufficient. From them, two more conditions can be deduced, namely:—

The vectors are all *independent*, i.e.

$$\text{if } \Sigma_r c_r \langle r| = 0, \quad \text{then } c_r = 0 \text{ for all } r, \tag{1·5}$$

and the operator

$$\Sigma_r |r\rangle \langle r| = 1. \tag{1·6}$$

To prove (1·5), multiply the equation $\Sigma_r c_r \langle r| = 0$ by $|s\rangle$ on the right, getting

$$\Sigma_r c_r \langle r \mid s \rangle = 0$$

or

$$c_s = 0$$

from (1·4). To prove (1·6), note that

$$\Sigma_r |r\rangle \langle r \mid s\rangle = \Sigma_r |r\rangle \delta_{rs} = |s\rangle.$$

Since the $|s\rangle$ form a complete set, like the $\langle r|$, we can infer that for any vector $|x\rangle$,

$$\Sigma_r |r\rangle \langle r \mid x\rangle = |x\rangle.$$

and (1·6) follows.

The conditions (1·5) and (1·6) are sufficient as well as necessary, since (1·3) and (1·4) can be deduced from them. To deduce (1·3) from them, take any vector $\langle x|$ and multiply it into (1·6) on the left. The result is

$$\Sigma_r \langle x \mid r\rangle \langle r| = \langle x|,$$

which gives $\langle x|$ expanded in the form (1·3), with $c_r = \langle x \mid r\rangle$. To deduce (1·4) from them, multiply (1·6) by $\langle s|$ on the left. The result is

$$\Sigma_r \langle s \mid r\rangle \langle r| = \langle s| = \Sigma_r \delta_{sr} \langle r|$$

or

$$\Sigma_r (\langle s \mid r\rangle - \delta_{sr}) \langle r| = 0.$$

Applying the independence condition (1·5), (1·4) follows.

The general formula (1·6) leads immediately to the law of matrix multiplication of quantum mechanics in all its various cases, e.g.

$$\langle s \mid \xi\eta \mid t\rangle = \Sigma_r \langle s \mid \xi \mid r\rangle \langle r \mid \eta \mid t\rangle$$

and

$$\langle r \mid \xi \mid x\rangle = \Sigma_s \langle r \mid \xi \mid s\rangle \langle s \mid x\rangle.$$

Instead of the discrete set of basic vectors $\langle r|$, one can have a continuous range of them, labelled by a parameter, q' say, that can take on all values in a certain continuous range. The sums in (1·3), (1·5), and (1·6) are then to be replaced by integrals over q', and (1·4) is to be replaced by

$$\langle q' \mid q''\rangle = \delta(q' - q''). \tag{1·7}$$

There now exists an operator, called the operator q, which, when operating on any $|\, q' \rangle$ or $\langle q' \,|$, has the effect of merely multiplying it by q', thus

$$q \,|\, q' \rangle = q' \,|\, q' \rangle \qquad \langle q' \,|\, q = q' \langle q' \,|. \tag{1·8}$$

There exists another important operator, which we may call the operator $\frac{d}{dq}$, which is such that when it operates on any $|\ \rangle$, the q' representative of that $|\ \rangle$ gets differentiated. In symbols

$$\langle q' \,|\, \frac{d}{dq} \,|\, \rangle = \frac{d}{dq'} \langle q' \,|\, \rangle.$$

Hence

$$\langle q' \,|\, \frac{d}{dq} = \frac{d}{dq'} \langle q' \,|\,, \tag{1·9}$$

which shows the effect of $\frac{d}{dq}$ operating on a basic $\langle\ |$. Now, with the help of (1·7),

$$\langle q' \,|\, \frac{d}{dq} \,|\, q'' \rangle = \frac{d}{dq'} \langle q' \,|\, q'' \rangle = \frac{d}{dq'} \delta (q' - q'') = - \frac{d}{dq''} \delta (q' - q'')$$

$$= - \frac{d}{dq''} \langle q' \,|\, q'' \rangle.$$

Hence

$$\frac{d}{dq} \,|\, q'' \rangle = - \frac{d}{dq''} \,|\, q'' \rangle, \tag{1·10}$$

showing the effect of $\frac{d}{dq}$ operating on a basis $|\ \rangle$.

We may check the commutation relation

$$\frac{d}{dq}\, q - q\, \frac{d}{dq} = 1. \tag{1·11}$$

We have

$$\langle q' \,|\, \frac{d}{dq}\, q = \frac{d}{dq'} \langle q' \,|\, q = \frac{d}{dq'} \{ q' \langle q' \,|\, \} = \langle q' \,|\, + q' \frac{d}{dq'} \langle q' \,|\,,$$

and again

$$\langle q' \,|\, q\, \frac{d}{dq} = q' \langle q' \,|\, \frac{d}{dq} = q' \frac{d}{dq'} \langle q' \,|.$$

Hence, subtracting

$$\langle q' \,|\, \left(\frac{d}{dq}\, q - q\, \frac{d}{dq} \right) = \langle q' \,|$$

and (1·11) follows.

One can have more general representations in which the basic vectors are labelled by several parameters $q_1', q_2' \ldots$ with continuous ranges of values. Corresponding to each of these parameters q_r' there is now an operator q_r and another operator $\frac{\partial}{\partial q_r}$.

With these representations, the representatives $\langle q_1' q_2' \ldots | \rangle$ of $|\rangle$ vectors are the wave functions of Schrödinger's theory. If such a wave function is normalized, it has the physical interpretation that $|\langle q_1' q_2' \ldots | \rangle|^2$ is the probability of the observables $q_1, q_2 \ldots$ having values close to $q_1', q_2' \ldots$ per unit range of variation of these values.

§ 2. *Fock's Treatment of the Harmonic Oscillator.*

A corner-stone of electrodynamic theory is the simple harmonic oscillator. It may be treated in quantum mechanics by Heisenberg's matrix method or by Schrödinger's wave functions, but a simpler method has been discovered by Fock,[2] which is much more suitable for our present purposes.

Neglecting unessential numerical coefficients, we have to deal with a coordinate q and momentum p satisfying

$$qp - pq = i . \tag{2·1}$$

Fock's method consists in introducing the variable

$$\xi = \frac{1}{\sqrt{2}} (p + iq) \tag{2·2}$$

and setting up a representation referring to ξ. Such a representation is of a more general kind than the representation referring to q of the preceding section, on account of ξ not being a real dynamical variable. We may take

$$\frac{d}{d\xi} = \frac{1}{\sqrt{2}} (p - iq) \tag{2·3}$$

since this leads to the commutation relation

$$\frac{d}{d\xi} \xi - \xi \frac{d}{d\xi} = \tfrac{1}{2}[(p - iq)(p + iq) - (p + iq)(p - iq)] = - i (qp - pq) = 1 ,$$

like (1·11).

The energy operator, excluding the zero point energy, is

$$\tfrac{1}{2}(p^2 + q^2) - \tfrac{1}{2} = \tfrac{1}{2}(p + iq)(p - iq) = \xi \frac{d}{d\xi} . \tag{2·4}$$

The eigenfunctions of this operator (which are the representatives of the eigen $|\rangle$ vectors) are ξ^n with $n = 0, 1, 2 \ldots$, giving the energy

[2] Fock, Zeits f. Phys. 49, p. 339 (1928). There is a correction in Phys. Zeits. de Soviet. U. 6, p. 428, footnote (1934).

levels. It follows that the representative of any $|\, a \,\rangle$ can be expressed as a power series

$$\langle\, \xi \,|\, a \,\rangle = a_0 + a_1 \xi + a_2 \xi^2 + a_3 \xi^3 \ldots \qquad (2{\cdot}5)$$

the various terms in the expansion corresponding to the various energy levels.

Let us determine the form of the representatives of the eigen $\langle\, |$ vectors. An eigen $\langle\, |$ satisfies

$$\langle\, |\, \xi \frac{d}{d\xi} = n \langle\, |$$

or

$$\langle\, |\, \xi \frac{d}{d\xi} \,|\, \xi' \,\rangle = n \langle\, |\, \xi' \,\rangle .$$

Assuming that relations like (1·8), (1·10) hold for the ξ-representation, we get

$$-\frac{d}{d\xi'} \{\xi' \langle\, |\, \xi' \,\rangle\} = n \langle\, |\, \xi' \,\rangle ,$$

showing that $\langle\, |\, \xi \,\rangle$ is of the form ξ^{-n-1}. It follows that the representative of any $\langle\, b\,|$ can be expressed as a power series

$$\langle\, b \,|\, \xi \,\rangle = b_0 \xi^{-1} + b_1 \xi^{-2} + b_2 \xi^{-3} + \ldots , \qquad (2{\cdot}6)$$

the various terms in the expansion again corresponding to the various energy levels.

One may look upon the representative $\langle\, \xi \,|\, a \,\rangle$ for any value of ξ as the scalar product of $|\, a \,\rangle$ with a basic $\langle\, \xi \,|$, and similarly $\langle\, b \,|\, \xi \,\rangle$ may be looked upon as the scalar product of $\langle\, b\,|$ with a basic $|\, \xi \,\rangle$. However, $\langle\, \xi\,|$ and $|\, \xi \,\rangle$ are not conjugate imaginary vectors. The rule that the same label is used for conjugate imaginary vectors does not apply when the label is a value of a complex dynamical variable, and the conjugate imaginary of $|\, \xi \,\rangle$ must be written $\langle\, \bar{\xi} \,|$.

Suppose that the $|\, a \,\rangle$ represented by the right-hand side of (2·5) and the $\langle\, b\,|$ represented by the right-hand side of (2·6) are conjugate imaginary vectors. Then the coefficients b_n and a_n, which refer to the same energy level, must be connected by a relation of the type

$$b_n = \lambda_n \bar{a}_n . \qquad (2{\cdot}7)$$

To determine λ_n we note that, if $\langle\, b\,|$ and $|\, a \,\rangle$ are conjugate imaginary vectors, then so are $\langle\, b \,|\, (p - iq)$ and $(p + iq) \,|\, a \,\rangle$ or $\langle\, b \,|\, \frac{d}{d\xi}$ and $\xi \,|\, a \,\rangle$. These vectors are represented by $\Sigma_0^\infty (n+1)\, b_n \xi^{-n-2}$ and $\Sigma_0^\infty a_n \xi^{n+1}$, which must therefore be connected in the same way in which the right-hand sides of (2·6) and (2·5) are connected by (2·7), so that

$$(n+1)\, b_n = \lambda_{n+1} \bar{a}_n .$$

Hence $$\lambda_{n+1} = (n+1)\lambda_n ,$$

and we may take

$$\lambda_n = n! .$$

The condition (2·7) now becomes

$$b_n = n!\,\bar{a}_n . \tag{2·8}$$

As a check, let us examine the requirement that $\langle b \mid (p+iq)$ and $(p-iq) \mid a \rangle$ shall also be conjugate imaginary vectors, or $\langle b \mid \xi$ and $\frac{d}{d\xi} \mid a \rangle$. The representatives of these vectors are $\sum_0^\infty b_n \xi^{-n}$ and $\sum_0^\infty n a_n \xi^{n-1}$. We see that these series are connected in the same way in which the right-hand sides of (2·6) and (2·5) are connected by (2·8), with the exception that there is a constant term b_0 in the former which is not connected with any term in the latter, and which has no analogue in the right-hand side of (2·6). To get over this difficulty, we must assume that an arbitrary constant may be added to the representative $\langle \mid \xi \rangle$ of any $\langle \mid$ and it still represents the same $\langle \mid$. Generalizing the argument, one finds the need for a more general assumption, that an arbitrary ascending power series $c_0 + c_1\xi + c_2\xi^2 + \ldots$ may be added to the representative $\langle \mid \xi \rangle$ of any $\langle \mid$ and it still represents the same $\langle \mid$.

The scalar product $\langle b \mid a \rangle$ of the vectors $\langle b \mid$ and $\mid a \rangle$ whose representatives are given by (2·6) and (2·5) must be of the form

$$\langle b \mid a \rangle = \mu_0 b_0 a_0 + \mu_1 b_1 a_1 + \mu_2 b_2 a_2 + \ldots ,$$

where the μ's are real positive numbers. By suitably redefining the coefficients a_n and b_n, one can arrange to have the μ's unity, so that

$$\langle b \mid a \rangle = b_0 a_0 + b_1 a_1 + b_2 a_2 + \ldots \tag{2·9}$$

This may be expressed as a contour integral round the origin,

$$\begin{aligned} \langle b \mid a \rangle &= \frac{1}{2\pi i} \oint (b_0 \xi^{-1} + b_1 \xi^{-2} + b_2 \xi^{-3} + \ldots)\, d\xi\, (a_0 + a_1\xi + a_2\xi^2 + \ldots) \\ &= \frac{1}{2\pi i} \oint \langle b \mid \xi \rangle d\xi \langle \xi \mid a \rangle . \end{aligned} \tag{2·10}$$

This result may be put into operator form

$$1 = \frac{1}{2\pi i} \oint \mid \xi \rangle d\xi \langle \xi \mid , \tag{2·11}$$

and is then the analogue of (1, 6) for the present kind of representation. We can now see how it comes about that an arbitrary ascending power series in ξ may be added to $\langle b \mid \xi \rangle$, since such a series would not affect the value of the integral (2·10).

Taking the vector $\langle b |$ to be the conjugate imaginary of $| a \rangle$, so that the labels b and a are the same and the coefficients b_r and a_r are connected by (2·8), we get from (2·9)

$$\langle a | a \rangle = \Sigma\, n! \,|\, a_n \,|^2 = \Sigma\, (n!)^{-1} \,|\, b_n \,|^2 . \qquad (2{\cdot}12)$$

These sums must be convergent for vectors that correspond to quantum states. It follows that the series $\Sigma a_n \xi^n$ representing $| a \rangle$ must be convergent for all ξ, but the series $\Sigma b_n \xi^{-n-1}$ representing $\langle b |$ need not converge for any ξ. For this reason one usually works only with $\langle \xi | \rangle$ representatives, and $\langle | \xi \rangle$ representatives are not suitable for general theoretical investigations.

The form of (2·12) shows that if $| a \rangle$ is normalized, it corresponds to a state for which the probability of the energy having the value n is

$$P_n = n! \,|\, a_n \,|^2 . \qquad (2{\cdot}13)$$

Let us work out the representatives of some simple operators. The unit operator is represented by

$$\langle \xi' | 1 | \xi'' \rangle = \langle \xi' | \xi'' \rangle = \frac{1}{\xi'' - \xi'} , \qquad (2{\cdot}14)$$

since this makes

$$\frac{1}{2\pi i} \oint \langle \xi' | 1 | \xi'' \rangle \, d\xi \, \langle \xi'' | a \rangle = \langle \xi' | a \rangle .$$

The operator ξ is represented by

$$\langle \xi' | \xi | \xi'' \rangle = \frac{\xi'}{\xi'' - \xi'} \quad \text{or} \quad \frac{\xi''}{\xi'' - \xi'} . \qquad (2{\cdot}15)$$

These two expressions count as equal in the representative of an operator, since their difference is unity, which does not contribute anything when multiplied into $\langle \xi'' | \rangle$ and integrated round the origin with respect to ξ'', and contributes only a constant when multiplied into $\langle | \xi' \rangle$ and integrated round the origin with respect to ξ'. The operator $\frac{d}{d\xi}$ is represented by

$$\langle \xi' | \frac{d}{d\xi} | \xi'' \rangle = \frac{1}{(\xi'' - \xi')^2} , \qquad (2{\cdot}16)$$

since this makes

$$\frac{1}{2\pi i} \oint \langle \xi' | \frac{d}{d\xi} | \xi'' \rangle \, d\xi'' \, \langle \xi'' | \rangle = \frac{d}{d\xi'} \langle \xi' | \rangle$$

$$\frac{1}{2\pi i} \oint \langle | \xi' \rangle \, d\xi' \, \langle \xi' | \frac{d}{d\xi} | \xi'' \rangle = - \frac{d}{d\xi''} \langle | \xi'' \rangle .$$

These results confirm the assumptions made in deriving (2·6).

The foregoing theory may be extended to several harmonic oscillators, a representation being set up with one complex variable ξ for each oscillator. A quantum state for a set of r oscillators is then represented as a power series in $\xi_1, \xi_2, \ldots \xi_r$, which may be written

$$\Sigma_n a_{n_1 n_2 \ldots n_r} \xi_1^{n_1} \xi_2^{n_2} \ldots \xi_r^{n_r},$$

summed for all $n_1, n_2, \ldots n_r$. If this function is properly normalized, the probability of the various oscillators being in the energy states $n_1, n_2, \ldots n_r$ respectively will be

$$P_{n_1 n_2 \ldots n_r} = n_1!\, n_2! \ldots n_r! \,|\, a_{n_1 n_2 \ldots n_r} \,|^2, \qquad (2{\cdot}17)$$

by an extension of (2·13).

§ 3. *Theory of Bosons by Fock's Method.*

The electromagnetic field consists of light-quanta or photons, which satisfy the Bose statistics. The general theory of bosons (particles satisfying Bose statistics) is thus needed as a basis for the quantum treatment of the electromagnetic field, and the best method of setting up this theory is one due to Fock.[3]

A state of an assembly of u bosons is represented by a symmetrical wave function $\langle q' q'' \ldots q^u \,|\, \rangle$, where each value for a variable q represents a state for one boson. Let us suppose for definiteness that the values a q can take on are discrete. Then $|\langle q' q'' \ldots q^u \,|\, \rangle|^2$ is the probability of the first particle being in the state q', the second in the state q'', and so on. Since, however, there is no means of distinguishing one particle from another, the only physically significant interpretation of the wave function is that, if $n^a, n^b, \ldots$ are the numbers of bosons in the various states $q^a, q^b, \ldots$ when the distribution $q', q'', \ldots, q^u$ occurs, so that $\Sigma n = u$, then

$$(u!/n^a!\, n^b! \ldots) |\langle q' q'' \ldots q^u \,|\, \rangle|^2 \qquad (3{\cdot}1)$$

is the total probability of there being these numbers of bosons in the various states.

If the number of bosons in the assembly is indeterminate, a state of the assembly must be represented by a series of component wave functions

$$\langle \cdot \,|\, \rangle, \quad \langle q' \,|\, \rangle, \quad \langle q' q'' \,|\, \rangle, \quad \langle q' q'' q''' \,|\, \rangle, \quad \ldots \qquad (3{\cdot}2)$$

referring successively to 0, 1, 2, 3, ... bosons, those referring to two or more bosons being symmetrical. Each component $\langle q' q'' \ldots q^u \,|\, \rangle$ determines the probability of there being u bosons distributed in any way over the

[3] Fock, Phys. Zeits. d. Soviet. U. 6 p. 425, § I (1934).

various states, in accordance with (3·1). $\Sigma_{q'\ldots q^u} |\langle q'\, q''\ldots q^u \mid \rangle|^2$ is the total probability of there being u bosons in the assembly, and this quantity summed for all u must equal unity.

Fock's method consists in introducing a set of variables $\xi_{q'}$, $\xi_{q''}$, . . ., one for each value of q or for each independent state of a boson, and representing a state of the assembly by the function $\langle \xi \mid \rangle$ of the ξ's

$$\langle \xi \mid \rangle = \Sigma_u\, u!^{-\frac{1}{2}} \Sigma_{q'\ldots q^u} \xi_{q'}\, \xi_{q''} \ldots \xi_{q^u} \langle q'\, q''\ldots q^u \mid \rangle . \quad (3{\cdot}3)$$

This function is a polynomial in the ξ's, the terms of any degree u corresponding to there being u bosons in the assembly. The terms of the form $\xi_{q^a}^{n^a}\, \xi_{q^b}^{n^b} \ldots$ correspond to there being n^a bosons in the state q^a, n^b in the state q^b, and so on, and hence the operator $\xi_{q^a}\, \partial / \partial\, \xi_{q^a}$ gives the number of bosons n^a in any state q^a. The coefficient of $\xi_{q^a}^{n^a}\, \xi_{q^b}^{n^b} \ldots$ equals the appropriate $\langle q'\, q''\ldots q^u \mid \rangle$ multiplied by $u!^{-\frac{1}{2}}\, (u!/n^a!\, n^b! \ldots)$, the factor in brackets ( ) being the number of equal terms occurring in the sum over all q's. Hence the probability (3·1) becomes

$$n^a!\, n^b! \ldots \mid \text{coefficient of } \xi_{q^a}^{n^a}\, \xi_{q^b}^{n^b} \ldots \text{ in } \langle \xi \mid \rangle \mid^2 . \quad (3{\cdot}4)$$

There arises a correspondence between the assembly of bosons and a set of harmonic oscillators, each of the boson states specified by a value of q corresponding to one oscillator. The power series in the ξ's representing a state of the assembly of bosons may be identified with that describing a state of the oscillators, the probability (3·4) of there being various numbers of bosons in the various states for a boson being the same as the probability (2·17) of the oscillators being in various quantum states. *The assembly of bosons and the set of oscillators are mathematically equivalent systems.*

The transformation from the set of components (3·2) to the function $\langle \xi \mid \rangle$ is a linear transformation and is therefore expressible in terms of a transformation function $\langle \xi \mid q'\, q''\ldots q^u \rangle$, thus

$$\langle \xi \mid \rangle = \Sigma_u\, \Sigma_{q'\ldots q^u} \langle \xi \mid q'\, q''\ldots q^u \rangle \langle q'\, q''\ldots q^u \mid \rangle . \quad (3{\cdot}5)$$

Comparing (3·5) with (3·3), one finds

$$\langle \xi \mid q'\, q''\ldots q^u \rangle = u!^{-\frac{1}{2}}\, \xi_{q'}\, \xi_{q''} \ldots \xi_{q^u} . \quad (3{\cdot}6)$$

The operators ξ_q, $\partial / \partial \xi_q$, when operating to the right, correspond to the emission of a boson into and the absorption of a boson from the state q respectively. From the theory of § 2, one sees that these operators are adjoint (i.e. $\langle a \mid \xi_q$ and $\partial / \partial \xi_q \mid a \rangle$ are conjugate imaginary vectors, and so are $\langle a \mid \partial / \partial \xi_q$ and $\xi_q \mid a \rangle$).

§ 4. *Relativistic Theory of a Particle.*

Let us take for simplicity a particle of zero rest-mass, as we shall afterwards apply the theory to photons. We want our work to be of relativistic form throughout, so we must first generalize the conditions of § 1 for basic vectors to get them into relativistic form.

A state of a particle without spin is represented by a wave function $\langle x_0 x_1 x_2 x_3 | \rangle$, or $\langle \mathbf{x} | \rangle$ for brevity, in which the time x_0 appears on the same footing as the spacial coordinates x_1, x_2, x_3. (The velocity of light is taken as unity.) The wave function satisfies the wave equation

$$\Box \langle \mathbf{x} | \rangle = 0 . \tag{4·1}$$

The value of the wave function for any point in space-time $\mathbf{x}$ may be regarded as the scalar product of a vector $| \rangle$ corresponding to the state and a basic vector $\langle \mathbf{x} |$. Since (4·1) always holds, the basic vectors $\langle \mathbf{x} |$ cannot be all independent, but must satisfy

$$\Box \langle \mathbf{x} | = 0 . \tag{4·2}$$

This marks a departure from (1·5). A similar equation must hold for the conjugate imaginary basic vectors

$$\Box | \mathbf{x} \rangle = 0 . \tag{4·3}$$

We need a rule for fixing the scalar product of a wave function $\langle \mathbf{x} | a \rangle$ and a conjugate wave function $\langle b | \mathbf{x} \rangle$, which rule has to be connected with the physical interpretation of the wave functions. The only Lorentz-invariant rule is the one put forward by Gordon and Klein, which involves abandoning (1·2) and thus has the disadvantage, from the physical side, of leading to negative probabilities in the interpretation, but which is quite satisfactory from the mathematical side, since (1·2) does not play any essential rôle in the mathematical development.

The Gordon-Klein rule makes

$$\langle b | a \rangle = \frac{1}{4\pi i} \iiint \left[\frac{\partial}{\partial x_0} \langle b | \mathbf{x} \rangle . \langle \mathbf{x} | a \rangle - \langle b | \mathbf{x} \rangle \frac{\partial}{\partial x_0} \langle \mathbf{x} | a \rangle \right] dx_1 \, dx_2 \, dx_3 ,$$

the integral being taken over any three-dimensional surface $x_0 =$ constant, and being easily seen to have the same value for all such surfaces. This result may be written for brevity

$$\langle b | a \rangle = (4\pi i)^{-1} \int \langle b | \mathbf{x} \rangle (\mathrm{I} - \Gamma) \langle \mathbf{x} | a \rangle \, dx , \tag{4·4}$$

where dx means $dx_1 \, dx_2 \, dx_3$ and I and Γ mean $\partial / \partial x_0$ operating on

the expression to the left and to the right respectively. It leads to the operator equation

$$1 = (4\pi i)^{-1} \int | \mathbf{x} \rangle (\mathrm{I} - \Gamma) \langle \mathbf{x} | \, dx \qquad (4{\cdot}5)$$

which replaces (1·6) in the present scheme.

As examples of (4·4), it follows that

$$\langle b | \mathbf{x}'' \rangle = (4\pi i)^{-1} \int \langle b | \mathbf{x}' \rangle (\mathrm{I}' - \Gamma') \langle \mathbf{x}' | \mathbf{x}'' \rangle \, dx$$

and

$$\langle \mathbf{x}' | a \rangle = (4\pi i)^{-1} \int \langle \mathbf{x}' | \mathbf{x}'' \rangle (\mathrm{I}'' - \Gamma'') \langle \mathbf{x}'' | a \rangle \, dx'' .$$

These results, considered as equations for the unknown function $\langle \mathbf{x}' | \mathbf{x}'' \rangle$, have the solution

$$\langle \mathbf{x}' | \mathbf{x}'' \rangle = -i\Delta (\mathbf{x}' - \mathbf{x}'') , \qquad (4{\cdot}6)$$

as is easily seen by taking the integrals in them to be over the three-dimensional surface $x_0' = x_0''$ and using the property of the Δ function given by Heisenberg and Pauli[4]

$$\frac{\partial \Delta (\mathbf{x})}{\partial x_0} = 4\pi \, \delta (x_1) \, \delta (x_2) \, \delta (x_3) \quad \text{for} \quad x_0 = 0 . \qquad (4{\cdot}7)$$

Equation (4·6) replaces (1·4) in the present scheme.

The relativistic quantum theory of a particle involves negative energies as well as negative probabilities. They may be examined by passing to the momentum-energy representation. On account of (4·2), the basic vectors $\langle \mathbf{x} |$ will have a Fourier resolution of the form

$$\langle \mathbf{x} | = (2\pi)^{-1} \Sigma_{k_0} \iiint e^{-i(\mathbf{k}, \mathbf{x})} \langle \mathbf{k} | \, k_0^{-1} \, dk_1 \, dk_2 \, dk_3 , \qquad (4{\cdot}8)$$

where $(\mathbf{k}, \mathbf{x})$ denotes the scalar product $k_0 x_0 - k_1 x_1 - k_2 x_2 - k_3 x_3$ and Σ_{k_0} denotes a sum over both values of k_0 that go with any k_1, k_2, k_3, namely

$$k_0 = \pm (k_1^2 + k_2^2 + k_3^2)^{\frac{1}{2}} . \qquad (4{\cdot}9)$$

The k_0^{-1} factor is inserted into the integrand of (4·8) since the combination $k_0^{-1} \, dk_1 \, dk_2 \, dk_3$ is Lorentz-invariant, so that the Fourier coefficient $\langle \mathbf{k} |$ defined by (4·8) is Lorentz-invariant. The vectors $\langle \mathbf{k} |$ thus introduced may be considered as the basic vectors of a representation, which is, apart from a factor $\hbar$, the momentum-energy representation, the energy p_0 and the momentum components p_1, p_2, p_3 being connected with $\mathbf{k}$ according to $p_\mu = k_\mu \hbar$.

---

[4] Heisenberg and Pauli. Zeits. f. Phys. 56, p. 1, equation (64) (1929). The Δ function we are using is minus that of Heisenberg and Pauli.

Put

$$k_0^{-1}\, dk_1\, dk_2\, dk_3 = \partial k$$

for brevity. The conjugate imaginary equation to (4·8) is

$$|\, \mathbf{x}\, \rangle = (2\pi)^{-1} \Sigma_{k_0} e^{i(\mathbf{k}, \mathbf{x})} |\, \mathbf{k}\, \rangle\, \partial k\,. \tag{4·10}$$

Thus

$$\frac{\partial |\, \mathbf{x}\, \rangle}{\partial x_0} \langle\, \mathbf{x}\, | - |\, \mathbf{x}\, \rangle \frac{\partial \langle\, \mathbf{x}\, |}{\partial x_0} = (2\pi)^{-2} i \Sigma_{k_0 k_0'} \iint e^{i(\mathbf{k}-\mathbf{k}', \mathbf{x})} (k_0 + k_0')\, |\, \mathbf{k}\, \rangle \langle\, \mathbf{k}'\, |\, \partial k\, \partial k'\,.$$

Integrating over all x_1, x_2, x_3 for constant x_0,

$$\int |\, \mathbf{x}\, \rangle (1 - \Gamma) \langle\, \mathbf{x}\, |\, dx = 2\pi i \Sigma_{k_0 k_0'}$$
$$\iint e^{i(k_0 - k_0')x_0}\, \delta(k_1 - k_1')\, \delta(k_2 - k_2')\, \delta(k_3 - k_3')(k_0 + k_0')\, |\, \mathbf{k}\, \rangle \langle\, \mathbf{k}'\, |\, \partial k\, \partial k'\,. \tag{4·11}$$

The integrand on the right vanishes unless $k_1 = k_1'$, $k_2 = k_2'$, $k_3 = k_3'$, which lead to $k_0 = \pm k_0'$. But it also vanishes if $k_0 = -k_0'$, and therefore vanishes unless $k_0 = k_0'$. Thus (4·11) reduces, with the help of (4·5), to

$$1 = \Sigma_{k_0} \int |\, \mathbf{k}\, \rangle \langle\, \mathbf{k}\, |\, \partial k\,. \tag{4·12}$$

This is the equation which replaces (1·6) for the **k**-representation. It is connected with the physical interpretation that the probability of k_1, k_2, k_3 having values within the ranges $dk_1\, dk_2\, dk_3$, with k_0 having either of the values (4·9), is

$$|\langle\, \mathbf{k}\, |\, a\, \rangle|^2\, \partial k = |\langle\, \mathbf{k}\, |\, a\, \rangle|^2\, k_0^{-1}\, dk_1\, dk_2\, dk_3\,, \tag{4·13}$$

for a state corresponding to a normalized vector $|\, a\, \rangle$. The probability (4·13) is positive or negative according to whether k_0 is positive or negative, which shows that states of positive energy always occur with a positive probability and those of negative energy with a negative probability.

From (4·12)

$$|\, \mathbf{k}'\, \rangle = \Sigma_{k_0} \int |\, \mathbf{k}\, \rangle \langle\, \mathbf{k}\, |\, \mathbf{k}'\, \rangle\, \partial k\,,$$

which shows, since the $|\, \mathbf{k}\, \rangle$ are all independent, that

$\langle\, \mathbf{k}\, |\, \mathbf{k}'\, \rangle = k_0\, \delta(k_1 - k_1')\, \delta(k_2 - k_2')\, \delta(k_3 - k_3')$ when k_0, k_0' have the same sign

$= 0$ when k_0, k_0' have opposite signs. (4.14)

This is the equation which replaces (1·4) for the **k**-representation.

§ 5. *Relativistic Theory of Bosons.*

The theory of bosons of § 3 may be made relativistic in a straightforward way. For bosons with no spin, one replaces each q by $\mathbf{x}$, denoting the four coordinates of the boson in space-time. Thus one gets, instead of (3·2) the series of component wave functions

$$\langle \cdot \mid \rangle, \ \langle \mathbf{x}' \mid \rangle, \ \langle \mathbf{x}' \mathbf{x}'' \mid \rangle, \ \langle \mathbf{x}' \mathbf{x}'' \mathbf{x}''' \mid \rangle, \ \ldots, \tag{5·1}$$

those beyond the second being symmetrical. Let us again take the case of particles of zero rest-mass. Then the various components (5·1) satisfy

$$\Box_{\mathbf{x}^r} \langle \mathbf{x}' \mathbf{x}'' \ldots \mathbf{x}^u \mid \rangle = 0, \quad \mathbf{x}^r = \mathbf{x}', \mathbf{x}'', \ldots, \mathbf{x}^u. \tag{5·2}$$

The Gordon-Klein rule can easily be extended to apply to a number u of bosons and reads, in the operator form corresponding to (4·5),

$$1 = (4\pi i)^{-u} \iint_{\ldots} |\mathbf{x}'\mathbf{x}'' \ldots \mathbf{x}^u\rangle (\mathrm{I}' - \Gamma')(\mathrm{I}'' - \Gamma'') \ldots (\mathrm{I}^u - \Gamma^u) \langle \mathbf{x}'\mathbf{x}'' \ldots \mathbf{x}^u | \, dx' dx'' \ldots dx^u. \tag{5·3}$$

The various operators I and Γ here all commute with each other, so their order does not matter. One can generalize (5·3) to apply to a variable number of bosons, by summing the right-hand side for all u.

One can alternatively work with the momentum representation or k-representation, replacing (5·1) by a set of functions referring to the k-variables

$$\langle \cdot \mid \rangle, \ \langle \mathbf{k}' \mid \rangle, \ \langle \mathbf{k}' \mathbf{k}'' \mid \rangle, \ \langle \mathbf{k}' \mathbf{k}'' \mathbf{k}''' \mid \rangle, \ \ldots$$

and one then has instead of (5·3), by extending (4·12),

$$1 = \Sigma_u \Sigma_{k_0' k_0'' \ldots k_0^u} \iint_{\ldots} |\mathbf{k}'\mathbf{k}'' \ldots \mathbf{k}^u\rangle \langle \mathbf{k}'\mathbf{k}'' \ldots \mathbf{k}^u | \, \partial k' \partial k'' \ldots \partial k^u. \tag{5·4}$$

We now introduce the ξ's, either $\xi_{\mathbf{x}}$, satisfying $\Box \xi_{\mathbf{x}} = 0$ or $\xi_{\mathbf{k}}$, defined for $k_0 = \pm (k_1^2 + k_2^2 + k_3^2)^{\frac{1}{2}}$. Corresponding to the non-relativistic theory of bosons, we represent a state of the assembly by a power series in the $\xi_{\mathbf{x}}$'s, using the transformation function

$$\langle \xi_{\mathbf{x}} \mid \mathbf{x}' \ \mathbf{x}'' \ldots \mathbf{x}^u \rangle = u!^{-\frac{1}{2}} \xi_{\mathbf{x}'} \xi_{\mathbf{x}''} \ldots \xi_{\mathbf{x}^u} \tag{5·5}$$

or as a power series in the $\xi_{\mathbf{k}}$'s, using the transformation function

$$\langle \xi_{\mathbf{k}} \mid \mathbf{k}' \ \mathbf{k}'' \ldots \mathbf{k}^u \rangle = u!^{-\frac{1}{2}} \xi_{\mathbf{k}'} \xi_{\mathbf{k}''} \ldots \xi_{\mathbf{k}^u}. \tag{5·6}$$

The first power series is, from (5·3)

$$\begin{aligned} \langle \xi_{\mathbf{x}} \rangle &= \Sigma_u (4\pi i)^{-u} \iint_{\ldots} \langle \xi_{\mathbf{x}} \mid \mathbf{x}' \mathbf{x}'' \ldots \mathbf{x}^u \rangle (\mathrm{I}' - \Gamma') \ldots \\ &\qquad (\mathrm{I}^u - \Gamma^u) \langle \mathbf{x}' \mathbf{x}'' \ldots \mathbf{x}^u \mid \rangle \, dx' \ldots dx^u \\ &= \Sigma_u (4\pi i)^{-u} u!^{-\frac{1}{2}} \iint_{\ldots} \xi_{\mathbf{x}'} \xi_{\mathbf{x}''} \ldots \xi_{\mathbf{x}^u} (\mathrm{I}' - \Gamma') \ldots \\ &\qquad (\mathrm{I}^u - \Gamma^u) \langle \mathbf{x}' \mathbf{x}'' \ldots \mathbf{x}^u \mid \rangle \, dx' \ldots dx^u \end{aligned} \tag{5·7}$$

and the second is similarly, from (5·4)

$$\langle \xi_{\mathbf{k}} | \rangle = \Sigma_u \, u!^{-\frac{1}{2}} \, \Sigma_{k_0' \ldots k_0^u} \int\int \ldots \xi_{\mathbf{k}'} \, \xi_{\mathbf{k}''} \ldots \xi_{\mathbf{k}^u}$$
$$\langle \mathbf{k}' \, \mathbf{k}'' \ldots \mathbf{k}^u | \rangle \, \partial k' \, \partial k'' \ldots \partial k^u . \qquad (5{\cdot}8)$$

The two power series are easily seen to be equivalent when one refers to the connection between $\xi_{\mathbf{x}}$ and $\xi_{\mathbf{k}}$

$$\xi_{\mathbf{x}} = (2\pi)^{-1} \, \Sigma_{k_0} \int e^{i(\mathbf{k}, \mathbf{x})} \, \xi_{\mathbf{k}} \, \partial k , \qquad (5{\cdot}9)$$

which corresponds to equation (4·10) for basic vectors. The coefficients in $\langle \xi_{\mathbf{x}} | \rangle$ may be used to give the probabilities of various bosons being in various places, and the coefficients in $\langle \xi_{\mathbf{k}} | \rangle$ may be used to give the probabilities of various bosons having various momenta. The latter of these probabilities are the more useful, so we shall consider them in more detail.

Let us replace the continuous range of k-states by a discrete set, by enclosing the bosons in a finite box, or otherwise, and let $s_{\mathbf{k}}$ be the density of the discrete set, per unit volume of (k_1, k_2, k_3)-space, in the neighbourhood of any $\mathbf{k}$. Integrals over k_1, k_2, k_3 are now to be replaced by sums, so that (5·8) becomes

$$\langle \xi_{\mathbf{k}} | \rangle = \Sigma_u \, u!^{-\frac{1}{2}} \, \Sigma_{\mathbf{k}' \ldots \mathbf{k}^u} \, \xi_{\mathbf{k}'} \, \xi_{\mathbf{k}''} \ldots \xi_{\mathbf{k}^u}$$
$$\langle \mathbf{k}' \, \mathbf{k}'' \ldots \mathbf{k}^u | \rangle \, (k_0' \, s_{\mathbf{k}'} \, k_0'' \, s_{\mathbf{k}''} \ldots k_0^u \, s_{\mathbf{k}^u})^{-1} .$$

From (4·13) transferred to the case of discrete k's, the probability of there being one boson and of its being in the state $\mathbf{k}$ is $|\langle k | \rangle|^2 \, (k_0 \, s_{\mathbf{k}})^{-1}$. Generalizing this result, we have that the probability of there being u bosons, n_a of them in state $\mathbf{k}^a$, n_b in state $\mathbf{k}^b$, and so on, is

$$(u!/n_a! \, n_b! \ldots) \, |\langle \mathbf{k}' \, \mathbf{k}'' \ldots \mathbf{k}^u | \rangle|^2 \, (k_0' \, s_{\mathbf{k}'} \, k_0'' \, s_{\mathbf{k}''} \ldots k_0^u \, s_{\mathbf{k}^u})^{-1} ,$$

where n_a of $\mathbf{k}' \, \mathbf{k}'' \ldots \mathbf{k}^u$ are equal to $\mathbf{k}^a$, n_b of them are equal to $\mathbf{k}^b$, and so on. This expression equals

$$n_a! \, n_b! \ldots \, | \text{ coefficient of } \xi_{\mathbf{k}^a}^{n_a} \, \xi_{\mathbf{k}^b}^{n_b} \ldots \text{ in } \langle \xi_{\mathbf{k}} | \rangle |^2$$
$$(k_0^a \, s_{\mathbf{k}^a})^{n_a} \, (k_0^b \, s_{\mathbf{k}^b})^{n_b} \ldots \qquad (5{\cdot}10)$$

For $k_0 > 0$, $\xi_{\mathbf{k}}$ may be considered to describe a harmonic oscillator. It must be taken as $(k_0 \, s_{\mathbf{k}})^{\frac{1}{2}}$ times the ξ of § 2 or 3, to get the same weight factor in (5·10) as in (3·4), and thus satisfies with its adjoint $\xi_{\mathbf{k}}^*$ the commutation relation

$$\xi_{\mathbf{k}}^* \, \xi_{\mathbf{k}} - \xi_{\mathbf{k}} \, \xi_{\mathbf{k}}^* = k_0 \, s_{\mathbf{k}} . \qquad (5{\cdot}11)$$

For $k_0 < 0$, one no longer has the analogy with the harmonic oscillator but (5·11) still holds, as may be verified thus :—

Consider the degree of freedom corresponding to $\xi_{\mathbf{k}}$ by itself, and take two states, a and b, represented by

$$\langle \xi_{\mathbf{k}} \mid a \rangle = \Sigma\, a_n\, \xi_{\mathbf{k}}^n$$
$$\langle \xi_{\mathbf{k}} \mid b \rangle = \Sigma\, b_n\, \xi_{\mathbf{k}}^n .$$

Corresponding to the physical interpretation (5·10), we must have the multiplication law

$$\langle a \mid b \rangle = \Sigma\, n!\, \bar{a}_n\, b_n\, (k_0\, s_{\mathbf{k}})^n .$$

Using this law

$$\langle a \mid \dot{\xi}_{\mathbf{k}} \mid b \rangle = \Sigma\, n!\, \bar{a}_n\, b_{n-1}\, (k_0\, s_{\mathbf{k}})^n .$$

Taking the conjugate complex equation,

$$\langle b \mid \xi_{\mathbf{k}}^* \mid a \rangle = \Sigma n!\, a_n \bar{b}_{n-1} (k_0 s_{\mathbf{k}})^n = \Sigma\, (n-1)!\, \bar{b}_{n-1} (n a_n k_0 s_{\mathbf{k}}) (k_0 s_{\mathbf{k}})^{n-1},$$

showing that $\xi_{\mathbf{k}}^* \mid a \rangle$ is represented by

$$\Sigma\, n a_n\, k_0\, s_{\mathbf{k}}\, \xi_{\mathbf{k}}^{n-1} = k_0\, s_{\mathbf{k}} \frac{d}{d\xi_{\mathbf{k}}} \langle \xi_{\mathbf{k}} \mid a \rangle .$$

Thus

$$\xi_{\mathbf{k}}^* = k_0\, s_{\mathbf{k}} \frac{d}{d\xi_{\mathbf{k}}}$$

and (5·11) follows, whether k_0 is positive or negative.

(5·11) leads to

$$\xi_{\mathbf{k}^a}^* \xi_{\mathbf{k}^b} - \xi_{\mathbf{k}^b} \xi_{\mathbf{k}^a}^* = k_0{}^a\, s_{\mathbf{k}^a}\, \delta_{ab} .$$

Passing back from discrete to continuous $\mathbf{k}$,

$$\left.\begin{aligned} \xi_{\mathbf{k}^a}^* \xi_{\mathbf{k}^b} - \xi_{\mathbf{k}^b} \xi_{\mathbf{k}^a}^* = k_0{}^a\, \delta(k_1{}^a - k_1{}^b)\, \delta(k_2{}^a - k_2{}^b)\, \delta(k_3{}^a - k_3{}^b) \\ \text{when } k_0{}^a,\ k_0{}^b \text{ have the same sign} \\ = 0 \quad \text{when } k_0{}^a,\ k_0{}^b \text{ have opposite signs.} \end{aligned}\right\} \quad (5\cdot12)$$

Returning to the $\xi_{\mathbf{x}}$'s, according to (5·9)

$$\xi_{\mathbf{x}'}^*\, \xi_{\mathbf{x}''} - \xi_{\mathbf{x}''}\, \xi_{\mathbf{x}'}^* = (2\pi)^{-2} \Sigma_{k'_0 k''_0} \iint e^{-i(\mathbf{k}',\, \mathbf{x}')}\, e^{i(\mathbf{k}'',\, \mathbf{x}'')}$$
$$(\xi_{\mathbf{k}'}^*\, \xi_{\mathbf{k}''} - \xi_{\mathbf{k}''}\, \xi_{\mathbf{k}'}^*)\, \partial k'\, \partial k''$$
$$= (2\pi)^{-2} \Sigma_{k'_0} \int e^{-i(\mathbf{k}',\, \mathbf{x}' - \mathbf{x}'')}\, \partial k'$$
$$= -2i(2\pi)^{-2} \int \sin(\mathbf{k}',\, \mathbf{x}' - \mathbf{x}'')\, \partial k',$$

where in the last expression k'_0 is restricted to be positive. Put $\mathbf{x}' - \mathbf{x}'' = \mathbf{x}$

and let $|x| = (x_1^2 + x_2^2 + x_3^2)^{\frac{1}{2}}$, and θ be the angle between the three-dimensional vectors x_1, x_2, x_3 and k_1, k_2, k_3. Then

$$\begin{aligned}
\xi^*_{\mathbf{x}'}\xi_{\mathbf{x}''} - \xi_{\mathbf{x}''}\xi^*_{\mathbf{x}'} &= -i\pi^{-1}\iint \sin[k_0(x_0 - |x|\cos\theta)]\, k_0\, dk_0\, d\cos\theta \\
&= -i\pi^{-1}|x|^{-1}\int_0^\infty \cos[k_0(x_0 - |x|\cos\theta)]_{\cos\theta=-1}^{\cos\theta=1}\, dk_0 \\
&= -i\pi^{-1}|x|^{-1}\int_0^\infty \{\cos[k_0(x_0 - |x|)] - \cos[k_0(x_0 + |x|)]\}\, dk_0 \\
&= -i|x|^{-1}\{\delta(x_0 - |x|) - \delta(x_0 + |x|)\} \\
&= -i\Delta(\mathbf{x}) = -i\Delta(\mathbf{x}' - \mathbf{x}'') . \qquad (5{\cdot}13)
\end{aligned}$$

The operators $\xi_{\mathbf{k}}$, $\xi^*_{\mathbf{k}}$, when operating to the right, correspond to the emission of a boson into and the absorption of a boson from the momentum-energy state $\mathbf{k}\hbar$, and likewise the operators $\xi_{\mathbf{x}}$, $\xi^*_{\mathbf{x}}$ correspond to the emission of a boson into and the absorption of a boson from the position state $\mathbf{x}$. The commutation relations (5·12), (5·13) which these operators satisfy should be compared with (4·14), (4·6) of the one-particle theory.

The above relativistic theory for spinless bosons of zero rest-mass may easily be extended to apply to photons. The wave function for a single boson must be made into a 4-vector with components corresponding to the four components of the electromagnetic potential. The 4-vector wave function may be considered as a function of the four coordinates $\mathbf{x}$ and of another variable μ taking on the four values 0, 1, 2, 3, and may be written $\langle \mathbf{x}_\mu | \rangle$. The μ here is not a suffix attached to $\mathbf{x}$, but is an independent variable, which is written in the lower position to express that the whole function $\langle \mathbf{x}_\mu | \rangle$ is subject to the contravariant law under Lorentz transformations, and which may be raised by the usual rule

$$\langle \mathbf{x}^\mu | \rangle = g^{\mu\nu}\langle \mathbf{x}_\nu | \rangle . \qquad (5{\cdot}14)$$

The multiplication rule (4·5) must be extended to

$$1 = -g^{\mu\nu}(4\pi i)^{-1}\int |\mathbf{x}_\mu\rangle(\Gamma - \Gamma)\langle \mathbf{x}_\nu| dx , \qquad (5{\cdot}15)$$

in which the $-$ sign is inserted with $g^{\mu\nu}$ for convenience, so as to give a factor $+1$ for μ and ν equal to 1, 2, or 3. Similarly, (4·6) must be extended to

$$\langle \mathbf{x}'_\mu | \mathbf{x}''_\nu \rangle = ig_{\mu\nu}\Delta(\mathbf{x}' - \mathbf{x}'') . \qquad (5{\cdot}16)$$

In a corresponding way the wave function for u particles must be made into a tensor of rank u and may be written $\langle \mathbf{x}'_\mu \mathbf{x}''_\nu \ldots \mathbf{x}^u_\tau | \rangle$, in which the $\mu, \nu, \ldots, \tau$ are independent variables from the $\mathbf{x}$'s. It must satisfy the

symmetry condition that it remains unchanged when any permutation is applied to the $\mathbf{x}$'s and the same permutation to $\mu, \nu, \ldots, \tau$. The multiplication rule (5·3) must be extended to

$$1 = (-4\pi i)^{-u} \iint\ldots |\mathbf{x}'_\mu \mathbf{x}''_\nu \ldots \mathbf{x}^{u}_\tau\rangle(1' - \Gamma')(1'' - \Gamma'')\ldots(1^u - \Gamma^u)$$
$$\langle \mathbf{x}'^\mu \mathbf{x}''^\nu \ldots \mathbf{x}^{u\tau}| \, dx' \, dx'' \ldots dx^u. \quad (5{\cdot}17)$$

The further development of these extensions is obvious and need not be mentioned in detail. The result is that the positional emission operator becomes a 4-vector $\xi_{\mathbf{x}\mu}$, and satisfies with its adjoint $\xi^*_{\mathbf{x}\mu}$ the commutation relation

$$\xi^*_{\mathbf{x}\mu}\xi_{\mathbf{x}'\nu} - \xi_{\mathbf{x}'\nu}\xi^*_{\mathbf{x}\mu} = ig_{\mu\nu}\Delta(\mathbf{x} - \mathbf{x}'), \quad (5{\cdot}18)$$

corresponding to (5·13), and similarly the momentum emission operator becomes a 4-vector $\xi_{\mathbf{k}\mu}$, satisfying with its adjoint $\xi^*_{\mathbf{k}\mu}$ the commutation relation

$$\left.\begin{aligned} \xi^*_{\mathbf{k}\mu}\xi_{\mathbf{k}'\nu} - \xi_{\mathbf{k}'\nu}\xi^*_{\mathbf{k}\mu} &= -g_{\mu\nu}k_0\,\delta(k_1 - k_1')\,\delta(k_2 - k_2')\,\delta(k_3 - k_3') \\ &\qquad \text{when } k_0, k_0' \text{ have the same sign} \\ &= 0 \qquad \text{when } k_0, k_0' \text{ have opposite signs} \end{aligned}\right\} \quad (5{\cdot}19)$$

corresponding to (5·12).

§ 6. *Classical Electrodynamics in Hamiltonian Form.*

For one electron, with coords z_μ and momenta p_μ, we have an elementary theory with the Hamiltonian

$$F \equiv -\frac{1}{2m}\left\{\left(\mathbf{p} - e\,\mathbf{A}(\mathbf{z})\right)^2 - m^2\right\} = 0, \quad (6{\cdot}1)$$

in which $\mathbf{A}(\mathbf{x})$ is a given field. The Poisson Bracket relations are

$$[p_\mu, z_\nu] = g_{\mu\nu}. \quad (6{\cdot}2)$$

The general equation of motion is

$$\frac{d\xi}{ds} = [\xi, F]. \quad (6{\cdot}3)$$

This gives

$$\frac{dz_\mu}{ds} = -\frac{\partial F}{\partial p^\mu} = \frac{1}{m}\left(p_\mu - e\,A_\mu(\mathbf{z})\right) \quad (6{\cdot}4)$$

the correct connection between velocity and momentum, and

$$\frac{dp_\mu}{ds} = \frac{\partial F}{\partial z^\mu} = \frac{e}{m}\left(p_\nu - eA_\nu(\mathbf{z})\right)\frac{\partial A^\nu(\mathbf{z})}{\partial z^\mu}$$

or

$$m\frac{d^2 z_\mu}{ds^2} = e\frac{dz^\nu}{ds}\left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}\right)_{\mathbf{z}}, \tag{6·5}$$

the correct equation of motion for an electron in the given field $\mathbf{A}(\mathbf{x})$.

We must generalise this to apply to several electrons and include the interaction between them. Let e_i, m_i, $z_{\mu i}$, $p_{\mu i}$, s_i refer to the ith electron. We now have several independent variables s_i, and a dynamical variable ξ is in general a function of all of them, satisfying equations of motion of the form

$$\frac{d\xi}{ds_i} = [\xi, F_i]. \tag{6·6}$$

For consistency of these equations of motion, we need $\dfrac{d^2\xi}{ds_i\,ds_j} = \dfrac{d^2\xi}{ds_j\,ds_i}$ for all ξ, or $[[\xi, F_j]F_i] = [[\xi, F_i], F_j]$.

From Poisson's identity $[\xi[F_i, F_j]] = 0$

or $[F_i, F_j] =$ a number.

Again, each F_i must equal zero throughout the motion, so we also need $\dfrac{dF_i}{ds_j} = 0$, or

$$[F_i, F_j] = 0, \tag{6·7}$$

which is a rather stronger condition than the above one.

Let us now assume:—

P.B. of two variables referring to different electrons $= 0$,

P.B. of $A_\mu(\mathbf{x})$ and any electron variable $= 0$,

F_i is of same form as for one electron,

$$F_i = -\frac{1}{2m_i}\left\{\left(\mathbf{p}_i - e_i\mathbf{A}(\mathbf{z}_i)\right)^2 - m_i^2\right\} = 0. \tag{6·8}$$

Then

$$\frac{dz_{\mu i}}{ds_j} = 0, \qquad \frac{dp_{\mu i}}{ds_j} = 0 \qquad \text{for } j \neq i, \tag{6·9}$$

showing that $z_{\mu i}$, $p_{\mu i}$, depend only on s_i, as they should. As for one electron

$$\frac{dz_{\mu i}}{ds_i} = \frac{1}{m_i}\left(p_{\mu i} - e_i A_\mu(\mathbf{z}_i)\right), \tag{6·10}$$

which is correct, and

$$m_i\frac{d^2 z_{\mu i}}{ds_i^2} = e_i\frac{dz^\nu_i}{ds_i}\left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}\right)_{\mathbf{z}_i}. \tag{6·11}$$

This must be made to agree with Lorentz's equation, including radiation damping, which is (putting $d\,\mathbf{z}_i/ds_i = \mathbf{v}_i$ and using dots for higher derivatives),

$$m_i \dot{v}_{\mu i} - \tfrac{2}{3} e_i^2 \ddot{v}_{\mu i} + \tfrac{2}{3} e_i^2 (\mathbf{v}\,\ddot{\mathbf{v}})\, v_{\mu i} = e_i v^{\nu}_i \{F_{\mu\nu\, in} + \Sigma_{j\neq i} F_{\mu\nu\, j\, ret}\} z_i \qquad (6{\cdot}12)$$

with

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial \alpha_\mu} - \frac{\partial A_\mu}{\partial x^\nu}\,. \qquad (6{\cdot}13)$$

Now it may be verified that on the world-line of an electron,[5]

$$F_{\mu\nu\, ret} - F_{\mu\nu\, adv} = \frac{4}{3} e\, (\ddot{v}_\mu v_\nu - \ddot{v}_\nu v_\mu)\,, \qquad (6{\cdot}14)$$

so (6·12) may be written

$$m_i \dot{v}_{\mu i} = e_i v^{\nu}_i \{F_{\mu\nu\, in} + \Sigma_{j\neq i} F_{\mu\nu\, j\, ret} + \tfrac{1}{2} F_{\mu\nu\, i\, ret} - \tfrac{1}{2} F_{\mu\nu\, i\, adv}\} z_i\,. \qquad (6{\cdot}15)$$

This agrees with (6·11) provided the $\mathbf{A}$ in (6·11) has the value

$$\mathbf{A}_{in} + \Sigma_{j\neq i} \mathbf{A}_{j\, ret} + \tfrac{1}{2} \mathbf{A}_{i\, ret} - \tfrac{1}{2} \mathbf{A}_{i\, adv} \qquad (6{\cdot}16)$$

when $\mathbf{x} = \mathbf{z}_i$.

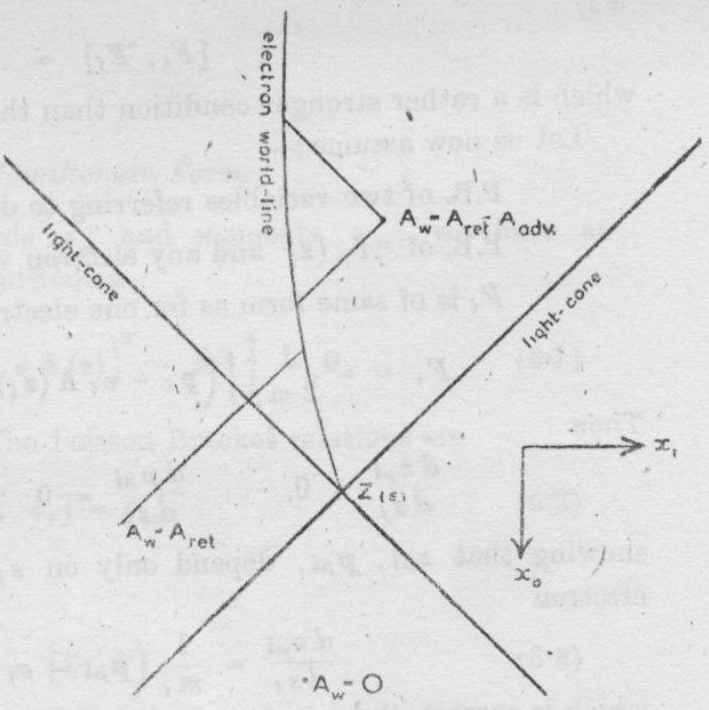
The *Wentzel field*,[6] $\mathbf{A}_W(\mathbf{x})$, is defined by

$$A_{\mu\, W}(\mathbf{x}) = A_{\mu\, in}(\mathbf{x}) + \Sigma_i e_i \int_{-\infty}^{s_i} v_{\mu i}(s'_i)\, \Delta\left(\mathbf{x} - \mathbf{z}_i(s'_i)\right) ds'_i\,. \qquad (6{\cdot}17)$$

It depends on the electron points s_i, which may be chosen anywhere on the world-lines of the electrons. The contribution of an electron to $\mathbf{A}_W$, namely

$$e \int_{-\infty}^{s} v'_\mu\, \Delta(\mathbf{x} - \mathbf{z}')\, ds'\,,$$

has the values shown in the figure.



[5] Dirac, Proc. Roy. Soc., A 167, p. 151 (1938).

[6] Wentzel, Zeits. f. Phys., 86, p. 479 (1933), gives the main properties of this field in the quantum theory.

$\mathbf{A}_W(\mathbf{x})$ has a singularity at each of the electron points $\mathbf{x} = \mathbf{z}_i$, its value at one of these points depending on the path along which one approaches the point. Suppose the electron points are all outside each other's light cones,

$$(\mathbf{z}_i - \mathbf{z}_j)^2 < 0 \qquad i \neq j . \tag{6·18}$$

Then $\mathbf{A}_W(\mathbf{x})$ has the value

(i) $\mathbf{A}_{in} + \Sigma_j \mathbf{A}_{j\,ret}$ if one approaches the point $\mathbf{x} = \mathbf{z}_i$ along a path outside the light-cone from $\mathbf{z}_i$,

(ii) $\mathbf{A}_{in} + \Sigma_{j\neq i} \mathbf{A}_{j\,ret}$ if one approaches the point $\mathbf{x} = \mathbf{z}_i$ along a path inside the future light-cone from $\mathbf{z}_i$,

(iii) $\mathbf{A}_{in} + \Sigma_j \mathbf{A}_{j\,ret} - \mathbf{A}_{i\,adv}$ if one approaches the point $\mathbf{x} = \mathbf{z}_i$ along a path inside the past light-cone from $\mathbf{z}_i$.

Expression (i) is just the Maxwell field. The mean of (ii) and (iii) is

$$\mathbf{A}_{in} + \Sigma_{j\neq i} \mathbf{A}_{j\,ret} + \tfrac{1}{2}\mathbf{A}_{i\,ret} - \tfrac{1}{2}\mathbf{A}_{i\,adv} ,$$

which is just (6·16). We can thus make (6·11) the correct equation of motion by taking

$$\mathbf{A}(\mathbf{x}) = \tfrac{1}{2}\{\mathbf{A}_W(\mathbf{x}+\boldsymbol{\lambda}) + \mathbf{A}_W(\mathbf{x}-\boldsymbol{\lambda})\} \tag{6·19}$$

where $\boldsymbol{\lambda}$ is a small time-like 4-vector, made to tend to zero ultimately, provided

$$(\mathbf{z}_i - \mathbf{z}_j \pm \boldsymbol{\lambda})^2 < 0 \qquad i \neq j . \tag{6·20}$$

This is a sharpening of condition (6·18). (6·20) may be looked upon as a restriction on the values $s_i, s_j \ldots$ of the independent variables. One may suppose that a dynamical variable is undefined except when this restriction is fulfilled.

It remains for us to get the correct equations of motion for the $\mathbf{A}_\mu(\mathbf{x})$. From (6·17)

$$\frac{dA_{\mu W}(\mathbf{x})}{ds_i} = e_i v_{\mu i} \Delta(\mathbf{x} - \mathbf{z}_i) ,$$

so that, with (6·19)

$$\frac{dA_\mu(\mathbf{x})}{ds_i} = \tfrac{1}{2} e_i v_{\mu i}\{\Delta(\mathbf{x} - \mathbf{z}_i + \boldsymbol{\lambda}) + \Delta(\mathbf{x} - \mathbf{z}_i - \boldsymbol{\lambda})\}. \tag{6·21}$$

The equation of motion

$$\frac{dA_\mu(\mathbf{x})}{ds_i} = [A_\mu(\mathbf{x}),\ F_i]$$

agrees with (6·21) provided we assume

$$[A_\mu(\mathbf{x}),\ A_\nu(\mathbf{x}')] = \tfrac{1}{2} g_{\mu\nu}\{\Delta(\mathbf{x} - \mathbf{x}' + \boldsymbol{\lambda}) + \Delta(\mathbf{x} - \mathbf{x}' - \boldsymbol{\lambda})\}. \tag{6·22}$$

This completes the scheme of P.B.'s and the Hamiltonian formulation. We have

$$[F_i, F_j] = 0$$

provided (6·20) holds, so the consistency condition (6·7) is fulfilled.

From (6·17)

$$\Box \mathbf{A}_W = 0 \tag{6·23}$$

since $\Box \Delta(\mathbf{x}) = 0$. Again

$$\begin{aligned}\frac{\partial}{\partial x_\mu} A_{\mu W} &= \Sigma_i e_i \int_{-\infty}^{s_i} v_{\mu i}' \frac{\partial}{\partial x_\mu} \Delta(\mathbf{x} - \mathbf{z}_i') \, ds_i' \\ &= -\Sigma_i e_i \int_{-\infty}^{s_i} v_{\mu i}' \frac{\partial}{\partial z_{\mu i}'} \Delta(\mathbf{x} - \mathbf{z}_i') \, ds_i' \\ &= -\Sigma_i e_i \int_{-\infty}^{s_i} \frac{d}{ds_i'} \Delta(\mathbf{x} - \mathbf{z}_i') \, ds_i' \\ &= -\Sigma_i e_i \Delta(\mathbf{x} - \mathbf{z}_i). \end{aligned} \tag{6·24}$$

These equations are to be contrasted with those for the Maxwell field $\mathbf{A}_M$

$$\Box \mathbf{A}_M \neq 0 \qquad \frac{\partial A_{\mu M}}{\partial x_\mu} = 0.$$

The Wentzel field can be resolved into waves travelling with the velocity of light.

Using (6·19)

$$\Box \mathbf{A} = 0, \tag{6·25}$$

$$\frac{\partial A_\mu(\mathbf{x})}{\partial x_\mu} = -\tfrac{1}{2}\Sigma_i e_i \{\Delta(\mathbf{x} - \mathbf{z}_i + \lambda) + \Delta(\mathbf{x} - \mathbf{z}_i - \lambda)\}. \tag{6·26}$$

These equations are not consequences of the equations of motion, and they therefore have to be imposed as extra conditions. One can verify that they are consistent with the equations of motion, *i.e.* if they hold initially they hold always. This follows since

$$[\Box \mathbf{A}, F_i] = 0 \tag{6·27}$$

and, putting

$$R(\mathbf{x}) = \frac{\partial A_\mu(\mathbf{x})}{\partial x_\mu} + \tfrac{1}{2}\Sigma_i e_i \{\Delta(\mathbf{x} - \mathbf{z}_i + \lambda) + \Delta(\mathbf{x} - \mathbf{z}_i - \lambda)\}, \tag{6·28}$$

we have

$$\begin{aligned}[R(\mathbf{x}), p_{\mu i} - e_i A_\mu(\mathbf{z}_i)] &= -\tfrac{1}{2} e_i \frac{\partial}{\partial z_{\mu_i}} \{\Delta(\mathbf{x} - \mathbf{z}_i + \lambda) + \Delta(\mathbf{x} - \mathbf{z}_i - \lambda)\} \\ &\quad - \tfrac{1}{2} e_i \frac{\partial}{\partial x_\mu} \{\Delta(\mathbf{x} - \mathbf{z}_i + \lambda) + \Delta(\mathbf{x} - \mathbf{z}_i - \lambda)\} \\ &= 0, \end{aligned} \tag{6·29}$$

so that

$$[R(\mathbf{x}), F_i] = 0. \tag{6·30}$$

Thus $\Box \mathbf{A}$ and $R(\mathbf{x})$ are in any case constants of the motion.

One can pass from the above many-time Hamiltonian formulation of the equations of motion to a one-time formulation, which is, of course, not of relativistic form, in the following way. First replace the Hamiltonian equations (6·8) by the equivalent equations (for $p_{0i} - e_i A_0 > 0$)

$$F_i' = -p_{0i} + e_i A_0(\mathbf{z}_i) + \{m_i^2 + \Sigma_r [p_{ri} - e_i A_r(\mathbf{z}_i)]^2\}^{\frac{1}{2}} = 0, \quad (6{\cdot}31)$$

where $r = 1$, 2, or 3. In terms of the F_i', the equations of motion are easily seen to be

$$\frac{d\xi}{dz_{0i}} = [\xi, F_i'].$$

Now put all the times z_{0i} equal to t. Then, for ξ independent of the p_{0i} and z_{0i},

$$\frac{d\xi}{dt} = [\xi, \Sigma_i F_i'(t)] = [\xi, H] \quad (6{\cdot}32)$$

where

$$H = \Sigma_i F_i'(t), \quad (6{\cdot}33)$$

or, as will do equally well,

$$H = \Sigma_i [e_i A_0 + \{m_i^2 + \Sigma_r (p_{ri} - e_i A_r)^2\}^{\frac{1}{2}}]. \quad (6{\cdot}34)$$

§ 7. *Elimination of the Longitudinal Waves.*

The field $A_\mu(\mathbf{x})$ may be split up into longitudinal and transverse waves, as follows :—

Split up the three-dimensional vector field $A_r(\mathbf{x})$, into a divergence-free part $M_r(\mathbf{x})$ and a rotation-free part $L_r(\mathbf{x})$, i.e.

$$A_r = M_r + L_r \quad (7{\cdot}1)$$

where

$$\Sigma_r \frac{\partial M_r}{\partial x_r} = 0 \quad (7{\cdot}2)$$

$$\frac{\partial L_r}{\partial x^s} - \frac{\partial L_s}{\partial x^r} = 0, \quad (7{\cdot}3)$$

and do this so that

$$\Box M_r = 0, \qquad \Box L_r = 0. \quad (7{\cdot}4)$$

Then M_r gives the transverse waves, and L_r and A_0 the longitudinal waves.

This splitting up is not relativistic, but it is of interest in spite of this, because the longitudinal waves can be eliminated from the Hamiltonian formulation by a contact transformation, thereby effecting a simplification of the equations.

The contact transformation can be given most conveniently by building it up from infinitesmal contact transformations, each dynamical variable ξ being transformed according to

$$\frac{d\xi^{(\tau)}}{d\tau} = [\xi^{(\tau)},\ G^{(\tau)}] \qquad 0 \leqslant \tau \leqslant 1 \qquad (7{\cdot}5)$$

from its initial value $\xi^{(0)} = \xi$ to its final value $\xi^{(1)} = \xi^*$ say.

Take
$$G^{(\tau)} = -\Sigma_j e_j \left[\int A_0^{(\tau)} dx_0\right]_{z_j}, \qquad (7{\cdot}6)$$

in which the arbitrary function of x_1, x_2, x_3 in $\int A_0^{(\tau)} dx_0$ is chosen so as to make

$$\Box \int A_0^{(\tau)} dx_0 = 0. \qquad (7{\cdot}7)$$

Then z_μ and $A_r(\mathbf{x})$ are invariant (i.e. independent of τ) and

$$\frac{dA_0(\mathbf{x})^{(\tau)}}{d\tau} = -\tfrac{1}{2}\Sigma_j e_j \left\{ \frac{1}{|x - z_j + \lambda|} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{|x - z_j - \lambda|} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} \qquad (7{\cdot}8)$$

where $|x|$ means the length of the 3-dimensional-vector x_r, and for the triplets $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ the first member is to be taken when $\mathbf{x}$ is in the past part of the light cone $\mathbf{z}_j - \lambda$ or $\mathbf{z}_j + \lambda$, the second member when it is outside the light cone, and the third member when it is the future part. Integrating (7·8)

$$A_0(\mathbf{x})^{(\tau)} = A_0(\mathbf{x}) - \tfrac{1}{2}\tau \Sigma_j e_j \left\{ \frac{1}{|x - z_j + \lambda|} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{|x - z_j - \lambda|} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} \qquad (7{\cdot}9)$$

and thus

$$A_0(\mathbf{z}_i)^{(\tau)} = A_0(\mathbf{z}_i) - \tfrac{1}{2}\tau \Sigma_{j \neq i} e_j \left\{ \frac{1}{|z_i - z_j + \lambda|} + \frac{1}{|z_i - z_j - \lambda|} \right\}. \qquad (7{\cdot}10)$$

Again
$$\frac{dp_{0i}^{(\tau)}}{d\tau} = -e_i A_0(\mathbf{z}_i)^{(\tau)}.$$

Using (7·10) and integrating,

$$p_{0i}^{(\tau)} = p_{0i} - \tau e_i A_0(\mathbf{z}_i) + \tfrac{1}{4}\tau^2 e_i \Sigma_{j \neq i} e_j \left\{ \frac{1}{|z_i - z_j + \lambda|} + \frac{1}{|z_i - z_j - \lambda|} \right\}.$$

Taking $\tau = 1$,

$$p_{0i} - e_i A_0(\mathbf{z}_i) = p_{0i}^* - \tfrac{1}{4} e_i \Sigma_{j \neq i} e_j \left\{ \frac{1}{|z_i - z_j + \lambda|} + \frac{1}{|z_i - z_j - \lambda|} \right\}. \qquad (7{\cdot}11)$$

Once more

$$\frac{dp_r^{(\tau)}}{d\tau} = -e_i \left[\int \frac{\partial A_0}{\partial x^r} dx_0 \right]_{\mathbf{z}_i}^{(\tau)}. \tag{7·12}$$

To evaluate this we need some subsidiary work. From (7·3)

$$L_r = \frac{\partial S}{\partial x^r}, \tag{7·13}$$

where S may be chosen so that

$$\Box S = 0. \tag{7·14}$$

Thus (6·26) gives, using (7·1) and (7·2),

$$\frac{\partial A_0}{\partial x_0} + \Sigma_r \frac{\partial^2 S}{\partial x_r \partial x^r} = -\tfrac{1}{2} \Sigma_i e_i \{\Delta(\mathbf{x} - \mathbf{z}_i + \boldsymbol{\lambda}) + \Delta(\mathbf{x} - \mathbf{z}_i - \boldsymbol{\lambda})\}$$

or

$$\frac{\partial A_0}{\partial x_0} - \frac{\partial^2 S}{\partial x_0^2} = -\tfrac{1}{2} \Sigma_i e_i \{\Delta(\mathbf{x} - \mathbf{z}_i + \boldsymbol{\lambda}) + \Delta(\mathbf{x} - \mathbf{z}_i - \boldsymbol{\lambda})\}.$$

Integrating with respect to x_0,

$$A_0 - \frac{\partial S}{\partial x_0} = \tfrac{1}{2} \Sigma_i e_i \left\{ \frac{1}{|x - z_i + \lambda|} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{|x - z_i - \lambda|} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} + f(x_1, x_2, x_3), \tag{7·15}$$

where the arbitrary function f must be chosen, from (7·14), so that $\Box$ applied to the right-hand side of (7·15) vanishes. From the theorem

$$\Box \frac{1}{|x|} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0 \tag{7·16}$$

which holds even at the origin, as may be verified by integrating (7·16) through a small volume around the origin and transforming to a surface integral, when one finds that it vanishes, one sees one must take $f = 0$.

Multiplying (7·15) by 2τ and adding to (7·9),

$$A_0^{(\tau)} - (1 - 2\tau) A_0 - 2\tau \frac{\partial S}{\partial x_0} =$$

$$= \tfrac{1}{2} \tau \Sigma_j e_j \left\{ \frac{1}{|x - z_j + \lambda|} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{|x - z_j - \lambda|} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Integrating with respect to x_0,

$$\int A_0^{(\tau)}\, dx_0 - (1-2\tau)\int A_0\, dx_0 - 2\tau S =$$

$$= \tfrac{1}{2}\tau\, \Sigma_j\, e_j \left\{ \frac{1}{|\, x - z_j + \lambda\,|} \begin{pmatrix} -|\, x - z_j + \lambda\,| \\ x_0 - z_{0j} + \lambda_0 \\ |\, x - z_j + \lambda\,| \end{pmatrix} + \right.$$

$$\left. + \frac{1}{|\, x - z_j - \lambda\,|} \begin{pmatrix} -|\, x - z_j - \lambda\,| \\ x_0 - z_{0j} - \lambda_0 \\ |\, x - z_j - \lambda\,| \end{pmatrix} \right\} + g(x_1, x_2, x_3), \qquad (7{\cdot}17)$$

where the arbitrary function g must be chosen, from (7·7) and (7·14), so that $\Box$ applied to the right-hand side of (7·17) vanishes. From the theorem

$$\Box \frac{1}{|\, x\,|} \begin{pmatrix} -|\, x\,| \\ x_0 \\ |\, x\,| \end{pmatrix} = 0, \qquad (7{\cdot}18)$$

which, like (7·16), holds even at the origin, one sees that $g = 0$. Using (7·17), (7·12) becomes

$$\frac{dp_{ri}^{(\tau)}}{d\tau} = -e_i(1-2\tau)\left[\int \frac{\partial A_0}{\partial x^r}\, dx_0\right]_{z_i} - 2\tau e_i \left(\frac{\partial S}{\partial x^r}\right)_{z_i} -$$

$$- \tfrac{1}{2}\tau e_i\, \Sigma_{j\neq i}\, e_j \left\{ \frac{(z_{ri} - z_{rj} + \lambda_r)(z_{0i} - z_{0j} + \lambda_0)}{|\, z_i - z_j + \lambda\,|^3} + \right.$$

$$\left. + \frac{(z_{ri} - z_{rj} - \lambda_r)(z_{0i} - z_{0j} - \lambda_0)}{|\, z_i - z_j - \lambda\,|^3} \right\}.$$

Integrating with respect to τ from 0 to 1,

$$p_{ri}^{*} - p_{ri} = -e_i \left(\frac{\partial S}{\partial x^r}\right)_{z_i} -$$

$$- \tfrac{1}{4} e_i\, \Sigma_{j\neq i}\, e_j \left\{ \frac{(z_{ri} - z_{rj} + \lambda_r)(z_{0i} - z_{0j} + \lambda_0)}{|\, z_i - z_j - \lambda\,|^3} + \right.$$

$$\left. + \frac{(z_{ri} - z_{rj} - \lambda_r)(z_{0i} - z_{0j} - \lambda_0)}{|\, z_i - z_j - \lambda\,|^3} \right\},$$

and thus

$$p_{ri} - e_i A_r(z_i) = p_{ri}^{*} - e_i M_r(z_i) +$$

$$+ \tfrac{1}{4} e_i\, \Sigma_{j\neq i}\, e_j \left\{ \frac{(z_{ri} - z_{rj} + \lambda_r)(z_{0i} - z_{0j} + \lambda_0)}{|\, z_i - z_j + \lambda\,|^3} + \right.$$

$$\left. + \frac{(z_{ri} - z_{rj} - \lambda_r)(z_{0i} - z_{0j} - \lambda_0)}{|\, z_i - z_j - \lambda\,|^3} \right\}. \qquad (7{\cdot}19)$$

This equation and (7·11) give expressions for $p_{\mu i} - e_i A_\mu(z_i)$ in terms of the new variables $p^*_{\mu i}$, $M_r(\mathbf{x})$, and by substituting these expressions into the Hamiltonians one eliminates the longitudinal waves.

In the limiting case $\lambda = 0$, (7·11) and (7·19) reduce to

$$\left.\begin{aligned} p_{0i} - e_i A_0(z_i) &= p^*_{0i} - \tfrac{1}{2} e_i \Sigma_{j \neq i} \frac{e_j}{|z_i - z_j|} \\ p_{ri} - e_i A_r(z_i) &= p^*_{ri} - e_i M_r(z_i) + \\ &\quad + \tfrac{1}{2} e_i \Sigma_{j \neq i} e_j \frac{(z_{ri} - z_{rj})(z_{0i} - z_{0j})}{|z_i - z_j|^3} . \end{aligned}\right\} \tag{7·20}$$

The one-time Hamiltonian (6·34) becomes, after elimination of the longitudinal waves, in the limit $\lambda = 0$,

$$H = \Sigma_i \Sigma_{j<i} \frac{e_i e_j}{|z_i - z_j|} + \Sigma_i \{ m_i^2 + \Sigma_r (p_{ri} - e_i M_r)^2 \}^{\frac{1}{2}} . \tag{7·21}$$

The ordinary Coulomb interaction energy appears here to replace the longitudinal waves.

§ 8. *Passage to the Quantum Theory.*

The preceding classical theory may be taken over into the quantum theory, by making the dynamical variables into operators satisfying commutation relations corresponding to the classical P.B. relations. No ambiguity concerning the order of factors arises. One may replace the form of Hamiltonian (6·1) for a spinless particle by the form

$$F = p_0 - e A_0(\mathbf{z}) - \Sigma_r \alpha^r (p_r - e A_r(\mathbf{z})) - \alpha_m m = 0 \tag{8·1}$$

for a spinning electron.

Each Hamiltonian F_i provides a wave equation

$$F_i | \rangle = 0. \tag{8·2}$$

To examine the consistency of these wave equations, note that they lead to

$$F_j F_i | \rangle = 0 \qquad F_i F_j | \rangle = 0$$

and hence

$$[F_i, F_j] | \rangle = 0 .$$

The equations are consistent since $[F_i, F_j] = 0$, for either spinless or spinning electrons, provided (6·20) holds.

The extra conditions $\Box \mathbf{A} = 0$ can be taken over unchanged into the quantum theory, since they are consistent with the commutation relations.

However, the extra conditions $R(\mathbf{x}) = 0$, with $R(\mathbf{x})$ defined by (6·28), are not consistent with the commutation relations, since they would lead to

$$[R(\mathbf{x}),\ A_\mu(\mathbf{x}')] = 0,$$

whereas actually

$$[R(\mathbf{x}),\ A_\mu(\mathbf{x}')] = \tfrac{1}{2}\frac{\partial}{\partial x^\mu}\{\Delta(\mathbf{x}-\mathbf{x}'+\lambda) + \Delta(\mathbf{x}-\mathbf{x}'-\lambda)\}.$$

The way out of the difficulty was shown by Fermi. It consists in assuming that the only states $|\ \rangle$ that occur in nature are those for which

$$R(\mathbf{x})\,|\ \rangle = 0. \tag{8·3}$$

There is one of these conditions for each $\mathbf{x}$, and we must see that they are all consistent with each other and with (8·2). We have

$$\begin{aligned}[R(\mathbf{x}),\ R(\mathbf{x}')] &= \left[\frac{\partial A_\mu(\mathbf{x})}{\partial x_\mu},\ \frac{\partial A_\nu(\mathbf{x}')}{\partial x'_\nu}\right] \\ &= \tfrac{1}{2}\frac{\partial^2}{\partial x_\mu\,\partial x'^\mu}\{\Delta(\mathbf{x}-\mathbf{x}'+\lambda) + \Delta(\mathbf{x}-\mathbf{x}'-\lambda)\} \\ &= -\tfrac{1}{2}\Box\{\Delta(\mathbf{x}-\mathbf{x}'+\lambda) + \Delta(\mathbf{x}-\mathbf{x}'-\lambda) = 0, \end{aligned} \tag{8·4}$$

showing that equations (8·3) are consistent with each other. From (6·29)

$$[R(\mathbf{x}),\ F_i] = 0,$$

whether F_i is of the form (8·1) or (6·1), and thus (8·3) is consistent with (8·2). Hence the conditions are consistent and we have a satisfactory mathematical scheme.

It remains for us to get a method of physical interpretation, by introducing a representation which will allow us to deduce from the wave functions the probabilities of various physical conditions holding. Since an interpretation of the electromagnetic field must involve photons, we try to connect the theory with the relativistic theory of photons of § 5. The obvious way of doing this is by putting

$$A_\mu(\mathbf{x}) = (\tfrac{1}{2}\hbar)^{\frac{1}{2}}(\xi_{\mathbf{x}-\frac{1}{2}\lambda,\,\mu} + \xi^*_{\mathbf{x}+\frac{1}{2}\lambda,\,\mu}), \tag{8·5}$$

since this gives from (5·18) the correct commutation relation for

$$[A_\mu(\mathbf{x}),\ A_\nu(\mathbf{x}')],$$

namely (6·22), and makes $A_\mu(\mathbf{x})$ self-adjoint in the limit $\lambda = 0$.

The wave function may now be taken to be a function of the coordinates $z_{\mu i}$ of the electrons and of the $\xi_{\mathbf{k}\mu}$, referring to the momenta of the photons. It is defined only for the z's satisfying (6·20) and it is a power

series, or rather a power integral in the $\xi_{\mathbf{k}\mu}$, containing, for example, linear terms like

$$\Sigma_{k_0'} \int c^{\mu}_{\mathbf{k}'} \, \xi_{\mathbf{k}'\mu} \, \partial k' \qquad (8{\cdot}6)$$

and quadratic terms like

$$\Sigma_{k_0' k_0''} \iint c^{\mu\nu}_{\mathbf{k}'\mathbf{k}''} \, \xi_{\mathbf{k}'\mu} \, \xi_{\mathbf{k}''\nu} \, \partial k' \, \partial k'' \qquad \left(c^{\mu\nu}_{\mathbf{k}'\mathbf{k}''} = c^{\nu\mu}_{\mathbf{k}''\mathbf{k}'} \right). \qquad (8{\cdot}7)$$

The variables $\xi^*_{k\mu}$ occurring in the Hamiltonians become operators of differentiation with respect to the $\xi_{\mathbf{k}\mu}$, with numerical coefficients chosen to give the correct commutation relations (5·19). Thus $\xi^*_{\mathbf{k}\lambda}$ applied to (8·6) produces $-c_{\mathbf{k}\lambda}$, and applied to (8·7) it produces

$$2 \Sigma_{k_0'} \int c^{\mu}_{\mathbf{k}'\mathbf{k}\lambda} \, \xi_{\mathbf{k}'\mu} \, \partial k'.$$

The coefficients in the power integral now give us the probability of the electrons being in specified places at specified times, with specified numbers of photons existing in the various momentum-energy states (in accordance with formula (5·10), if one passes from continuous to discrete **k**-values). This method of interpretation involves negative photon energies and negative probabilities. The way to understand these will be discussed later.

Longitudinal photons have no practical significance. It is therefore desirable to eliminate the longitudinal electromagnetic waves, which can be done by taking over the work of § 7 into the quantum theory. (This work was arranged so that every step in it can be immediately transferred to the quantum theory.) One then needs to use (8·5) only to connect the transverse part of $A_\mu(\mathbf{x})$, namely $M_\mu(\mathbf{x})$, with the transverse parts of $\xi_{\mathbf{x}\mu}$ and $\xi^*_{\mathbf{x}\mu}$, and only the transverse components of $\xi_{\mathbf{k}\mu}$ will appear in the wave function, i.e. those linear combinations of $\xi_{\mathbf{k}_1}$, $\xi_{\mathbf{k}_2}$, $\xi_{\mathbf{k}_3}$ which correspond to directions in three-dimensional space perpendicular to k_1, k_2, k_3.

After eliminating the longitudinal waves, it becomes practicable to solve the wave equations by a perturbation method, treating those terms in the Hamiltonians that contain field variables as the perturbation. The important question now arises, whether the integrals arising in the solution are all convergent for high **k**-values. If they are not, the whole theory would be useless.

From (8·5) and (5·9) one sees that $\xi_{\mathbf{k}\mu}$, $\xi^*_{\mathbf{k}\mu}$ occur in the Hamiltonians only in the combinations

$$\Sigma_{k_0} \int e^{i(\mathbf{k}, \mathbf{z}_i - \frac{1}{2}\lambda)} \, \xi_{\mathbf{k}\mu} \, \partial k, \qquad \Sigma_{k_0} \int e^{-i(\mathbf{k}, \mathbf{z}_i + \frac{1}{2}\lambda)} \, \xi^*_{\mathbf{k}\mu} \, \partial k. \qquad (8{\cdot}8)$$

With the perturbation method of solving the wave equations, the various terms in the wave function are given by repeated applications to the original wave function of the operators (8·8) and of other operators involving the electron momentum variables in a rational algebraic way and causing the appearance of rational algebraic functions of the k's. An application of the first type of operator (8·8) does not lead to a numerical integral over the k's, but an application of the second type leads to a numerical integral of the form

$$\Sigma_{k_0}\int f(\mathbf{k})\, e^{-i(\mathbf{k},\,\lambda)}\,\partial k \quad \text{or} \quad \Sigma_{k_0}\int f(\mathbf{k})\, e^{i(\mathbf{k},\,\mathbf{z}_i-\mathbf{z}_j-\lambda)}\,\partial k \qquad (8{\cdot}9)$$

with $f(\mathbf{k})$ rational algebraic. Carrying out the integration first with respect to k_0, taking fixed ratios $k_0 : k_1 : k_2 : k_3$, one has an integral of the form

$$\int_{-\infty}^{\infty} g(k_0)\, e^{-ik_0 a}\, dk_0\,, \qquad (8{\cdot}10)$$

where a is a number of the order λ or $\mathbf{z}_i - \mathbf{z}_j$, and $g(k_0)$ is rational algebraic. Thus for $|k_0|$ very large, $g(k_0)$ is of the form k_0^{2n} or k_0^{2n+1}. It is reasonable on physical grounds to approach the upper and lower limits of integration in (8·10) at the same rate, so the integral becomes of the form

$$\int^{\infty} k_0^{2n} \cos k_0 a\, dk_0 \quad \text{or} \quad \int^{\infty} k_0^{2n+1} \sin k_0 a\, dk_0\,. \qquad (8{\cdot}11)$$

Such integrals oscillate at the upper limit, but it is reasonable to take the mean of the oscillating part (as one does in other cases in quantum mechanics, e.g. in evaluating $\langle q' \mid p^n \mid q'' \rangle = h^{-\frac{1}{2}}\int_{-\infty}^{\infty} p^n\, e^{ip(q'-q'')/\hbar}\, dp$) and one then gets a definite value, which remains finite as $\lambda \to 0$, causing possibly $a \to 0$.

It is necessary that both positive and negative values of k_0 should occur in (8·8) and (8·9), otherwise one would have, as well as (8·11), integrals of the form

$$\int^{\infty} k_0^{2n+1} \cos k_0 a\, dk_0\,, \quad \int^{\infty} k_0^{2n} \sin k_0 a\, dk_0\,, \qquad (8{\cdot}12)$$

and no procedure of taking the mean of an oscillating part would enable one to assign a finite value to such integrals in the limit $a = 0$. One would get instead terms of the form a^{-2n-2}, a^{-2n-1}. The earlier quantum electrodynamics of Heisenberg and Pauli, which worked with only positive-energy photons, was faced with just this difficulty.

§ 9. *The Redundant Variables.*

Corresponding to (8·5), define

$$B_\mu(\mathbf{x}) = (\tfrac{1}{2}\hbar)^{\frac{1}{2}} (\xi_{\mathbf{x}+\frac{1}{2}\lambda,\,\mu} - \xi^*_{\mathbf{x}-\frac{1}{2}\lambda,\,\mu}), \tag{9·1}$$

which provides us with a new field. We have, from (5·18),

$$[B_\mu(\mathbf{x}),\ A_\nu(\mathbf{x}')] = 0,$$

and thus the $B_\mu(\mathbf{x})$ commute with all the dynamical variables occurring in the Hamiltonians. Such quantities are called redundant variables. They are, of course constants of the motion. If the longitudinal part of $A_\mu(\mathbf{x})$ has been eliminated, then the transverse part of $B_\mu(\mathbf{x})$ provides the only interesting redundant variables.

Let us examine the significance of redundant variables in an elementary case. A system of one degree of freedom, with the canonical variables q, p, may be considered as a system of two degrees of freedom by putting

$$q = \frac{1}{\sqrt{2}}(q_1 + q_2), \quad p = \frac{1}{\sqrt{2}}(p_1 + p_2). \tag{9·2}$$

Then $q_1 - q_2$ and $p_1 - p_2$ are redundant variables, as they commute with q and p. Make a canonical transformation which separates the redundant and non-redundant variables, by putting

$$q_r = \frac{1}{\sqrt{2}}(q_1 - q_2), \quad p_r = \frac{1}{\sqrt{2}}(p_1 - p_2), \tag{9·3}$$

so that (9·2) and (9·3) give the required contact transformation. Referred to the new variables, the wave function $\langle qq_r | \rangle$ is a function of q, q_r satisfying a wave equation which involves only q and $\frac{\partial}{\partial q}$. It must therefore satisfy this wave equation for each value of q_r, which thus appears in it only as a parameter. Since q_r has no physical meaning, to get a physical interpretation we must integrate $|\langle qq_r | \rangle|^2$ for all q_r, and so get the probability of q having a specified value or lying in a specified range.

One can now see that $\langle qq_r | \rangle$ does not represent a pure quantum state. Such a state is represented by a function of the single variable q satisfying the wave equation. (It is the state ensuing from a maximum observation, in this case an observation of q.) $\langle qq_r | \rangle$ represents a mixture of several pure states, namely, all those pure states represented by functions of q, which one can get by giving numerical values to q_r in $\langle qq_r | \rangle$. Such a mixture is the analogue of a Gibbs ensemble in classical mechanics. One can conclude that:—*The existence of redundant variables means that a wave function represents a Gibbs ensemble and not a pure state.* This conclusion must still hold if one is using a representation which does not separate the redundant and non-redundant variables.

One could try to make a canonical transformation to separate the redundant and non-redundant variables in our present quantum electrodynamics. Let us work with the limiting case $\lambda = 0$. Put

$$\xi_{-k\mu} = \zeta_{k\mu} \quad \text{for} \quad k_0 > 0. \tag{9·4}$$

Then, from (5·9), (8·5), and (9·1),

$$A_\mu(\mathbf{x}) = (\tfrac{1}{2}\hbar)^{\frac{1}{2}} (2\pi)^{-1} \iiint \{(\xi_{k\mu} - \zeta^*_{k\mu}) e^{i(kx)} + (-\zeta_{k\mu} + \xi^*_{k\mu}) e^{-i(kx)}\} \partial k \tag{9·5}$$

$$B_\mu(\mathbf{x}) = (\tfrac{1}{2}\hbar)^{\frac{1}{2}} (2\pi)^{-1} \iiint \{(\xi_{k\mu} + \zeta^*_{k\mu}) e^{i(kx)} + (-\zeta_{k\mu} - \xi^*_{k\mu}) e^{-i(kx)}\} \partial k \tag{9·6}$$

with $k_0 = +(k_1^2 + k_2^2 + k_3^2)^{\frac{1}{2}}$. Thus putting

$$\left.\begin{aligned} \sqrt{2}a_{k\mu} &= \xi_{k\mu} - \zeta^*_{k\mu} & \sqrt{2}a^*_{k\mu} &= \xi^*_{k\mu} - \zeta_{k\mu} \\ \sqrt{2}\beta_{k\mu} &= \zeta_{k\mu} + \xi^*_{k\mu} & \sqrt{2}\beta^*_{k\mu} &= \zeta^*_{k\mu} + \xi_{k\mu} \end{aligned}\right\}, \tag{9·7}$$

a and a^* are the only field variables occurring in the Hamiltonians, and β and β^* commute with a and a^* and are the redundant variables.

Let us confine our attention to one particular k-value (with $k_0 > 0$), which we may suppose to be one of a discreet set, and one particular μ-value, corresponding to transverse waves. Then, using (5·19) and dropping the suffixes,

$$\left.\begin{aligned} \xi^*\xi - \xi\xi^* &= c \\ \zeta^*\zeta - \zeta\zeta^* &= -c \end{aligned}\right\}, \tag{9·8}$$

where c is some positive number. From (9·7) and (9·8)

$$\left.\begin{aligned} a^*a - aa^* &= c \\ \beta^*\beta - \beta\beta^* &= -c. \end{aligned}\right\} \tag{9·9}$$

One can now set up a contact transformation from the variables ξ, ξ^*, ζ, ζ^*, to the variables a, a^*, β, β^*. The transformation function has been worked out by the author.[7] It enables one to transform the wave function from a power series in ξ, ζ to a power series in a, β. [From the commutation relations (9·8), (9·9), two forms of transformation are possible, one of which transforms an ascending power series in ξ, ζ to an ascending power series in a, β, the other transforms it to a descending power series in a, β. We choose the former, which is the one worked out in the reference.[7]]

The Hamiltonians get transformed to functions of a, $\partial/\partial a$ independent of β, $\partial/\partial\beta$. Thus the redundant variable β may be eliminated from the

[7] Dirac, Proc. Roy. Soc. A. 180, p. 1 (1942). Appendix III.

wave equations. However, if one carries out this elimination procedure for all k-values simultaneously, one gets the wave equations of Heisenberg and Pauli's quantum electrodynamics, which do not have any solution, because of the divergent integrals discussed at the end of § 8. Thus one cannot apply the transformation which eliminates all the redundant variables to solutions of the wave equations. All the same this transformation has a meaning because one can apply it to suitable functions of the ξ's, ζ's not satisfying the wave equations.

§ 10. *The Negative Energies and Probabilities.*

We may give the redundant variables any values we like, subject to no inconsistency arising from their Poisson Bracket relationships. It leads to a reasonable interpretation for the negative energies and probabilities to take the transverse part of $B_\mu(x)$ equal to the transverse part of the initial value of $A_\mu(x)$, after elimination of the longitudinal waves. With this assumption, for the initial state $\mid init \rangle$,

$$[A_\mu(\mathbf{x})_{transverse} - B_\mu(\mathbf{x})_{transverse}] \mid init \rangle = 0. \qquad (10{\cdot}1)$$

Making the Fourier resolution of this equation we get, from (9·5) and (9·6),

$$\xi^*_{\mathbf{k}t} \mid init \rangle = 0 \qquad \zeta^*_{\mathbf{k}t} \mid init \rangle = 0, \qquad (10{\cdot}2)$$

where t denotes a transverse component. These conditions are evidently all consistent, since the ξ^*'s and ζ^*'s all commute. They show that

$$\frac{\partial}{\partial \xi_{\mathbf{k}t}} \langle \xi\zeta \mid init \rangle = 0 \qquad \frac{\partial}{\partial \zeta_{\mathbf{k}t}} \langle \xi\zeta \mid init \rangle = 0, \quad (10{\cdot}3)$$

so that the representative $\langle \xi\zeta \mid init \rangle$ of the initial state is independent of the ξ's and ζ's, and is a function only of the electron variables $z_{\mu i}$. It thus corresponds to no photons being present, of either positive or negative energies.

The following natural interpretation for the probabilities (5·10) given by the wave function at some later time now appears. That part of the wave function corresponding to no photons present may be supposed to give the probability of no change having taken place in the field of photons; that part corresponding to one positive-energy photon present may be supposed to give the probability of a photon having been emitted; that corresponding to one negative-energy photon present may be supposed to give the probability of a photon having been absorbed; and so on for the parts corresponding to two or more photons present. *The various parts of the wave function which referred*

to the existence of positive- and negative-energy photons in the old interpretation now refer to the emissions and absorptions of photons. This disposes of the negative-energy difficulty in a satisfactory way, conforming to the laws of conservation of energy and momentum. It is possible only because of the redundant variables enabling one to arrange that the initial wave function shall correspond in its entirety to no emissions or absorptions having taken place.

The assumptions (10·1) or (10·2) are not compatible with the equations of motion and can therefore hold only for a particular value for the particle times z_{0i}. These assumptions are thus suitable only for problems in which the initial conditions apply to one time for each particle, for example if the initial positions of the particles in space-time are given. In practical problems one is usually given, not the initial positions of the particles, but their initial momenta for some collision process. The momenta $p_{\mu i}$ do not commute with the Hamiltonians (6·1) or (8·1) and the requirement that they should have certain initial values means that one must solve the wave equations by a perturbation method, taking as the zero-order wave function one for which they have these values and arranging that the higher-order wave-functions should refer only to outgoing particles or else latent particles (i.e. particles in transient states for which the momenta have values not satisfying $p_{\mu i}\, p^{\mu}{}_{i} = m_i{}^2$). The zero-order wave function in this perturbation method is then the initial wave function for which the conditions (10·1) or (10·2) hold, and any emission or absorption of photons that takes place according to the new interpretation given above refers to changes from the state represented by this zero-order wave function.

The physical interpretation is not yet complete, because at present it would give a negative probability for a process involving the absorption of a photon, or the absorption of any odd number of photons. To track down these negative probabilities, let us study the initial state by transforming the initial wave function to the α, β variables introduced in § 9. From (10·2) and (9·7)

$$(\alpha^* + \beta) \mid init \rangle = 0\,, \qquad (\beta^* - \alpha) \mid init \rangle = 0\,, \tag{10·4}$$

which give, from the commutation relations (9·9)

$$\left(c\frac{\partial}{\partial \alpha} + \beta\right) \langle \alpha\beta \mid init \rangle = 0 \qquad \left(c\frac{\partial}{\partial \beta} + \alpha\right) \langle \alpha\beta \mid init \rangle = 0\,. \tag{10·5}$$

Hence
$$\langle \alpha\beta \mid init \rangle = constant\ e^{-\alpha\beta/c}\,. \tag{10·6}$$

Thus, applying (5·10) to the α, β variables, the probability of there being for the initial state m photons of the **A** field and n of the **B** field in the

photon state considered is zero if $m \neq n$ and is proportional to $(-1)^n$ if $m = n$. The probability of there being n photons in the **A** field and in the **B** field must be taken equal to

$$P_n = 2(\varepsilon - 1)^n, \tag{10·7}$$

where ε is a small positive quantity tending to zero, in order that we may have $\Sigma_0^\infty P_n = 1$.

The probabilities (10·7), equal to 2 and -2, are not physically understandable, but one can use them mathematically in accordance with the rules for working with a Gibbs ensemble. One can suppose a hypothetical mathematical world with the initial probability distribution (10·7) for the photons, and one can work out the probabilities of radiative transition processes occurring in this world. One can deduce the corresponding probability coefficients, i.e. the probabilities per unit intensity of each beam of incident radiation concerned, by using Einstein's laws of radiation. For example, for a process involving the absorption of a photon, if the probability coefficient is B, the probability of the process is

$$\Sigma_{n=0}^\infty nP_n B = -\tfrac{1}{2}B, \tag{10·8}$$

and for a process involving the emission of a photon, if the probability coefficient is A, the probability of the process is

$$\Sigma_{n=0}^\infty (n+1) P_n A = \tfrac{1}{2}A. \tag{10·9}$$

Now the probability of an absorption process, as calculated from the theory, is negative, and that for an emission process is positive, so that, equating these calculated probabilities to (10·8) and (10·9) respectively, one obtains positive values for both B and A. Generally, it is easily verified that any radiative transition probability coefficient obtained by this method is positive.

It now becomes reasonable to assume that *these probability coefficients obtained for a hypothetical world are the same as those of the actual world.* One gets in this way a general physical interpretation for the quantum theory of photons. When applied to elementary examples, it gives the same results as Heisenberg and Pauli's quantum electrodynamics with neglect of the divergent integrals, since the extra factor $2^{-\frac{1}{2}}$ occurring in the matrix elements of the present theory owing to the $\sqrt{2}$ in (9·7) compensates the factor $\frac{1}{2}$ in the right-hand side of (10·8) or (10·9).

The above interpretation enables one to calculate the probability coefficients for all transition processes, which is essentially all that one needs for comparing the theory with experiment. It involves the logical

defect that it requires one to assume Einstein's laws of radiation, instead of enabling one to deduce them from the theory. This defect can be reduced, however, by the circumstance that one can deduce Einstein's laws of radiation for the hypothetical world by working with initial wave functions which are not completely independent of the ξ's and ζ's, as (10·3) requires, but contain small powers of some of the ξ's and ζ's, and then seeing how the resulting change in the initial probability distribution of the photons (10·7) affects the probability for transition processes.[8] It then remains only to assume that the same laws of radiation hold for the actual world as for the hypothetical one, an assumption of the same nature as the main assumption above.

A further point in connection with the interpretation should be noted. The theory involves processes in which a certain photon is emitted and the same photon is absorbed—two actions which cancel each other and leave nothing observable. However, according to Einstein's laws, such processes would be stimulated by incident radiation in a different way from what they would be if these actions did not occur, and thus there is a possibility of getting experimental evidence for such actions.

[8] This point was brought up by correspondence with W. Pauli.